RR

Vol. 47, Parts 1 and 2 June 1960

UNIVERSITY OF MICHIGAN
JUL 19 1960
SCIENCE LIBRARY

Science
QH
301
.B615
v.47
no 1/2

BIOMETRIKA

FOUNDED BY

W. F. R. WELDON, FRANCIS GALTON AND KARL PEARSON

MANAGING EDITOR

E. S. PEARSON

ASSOCIATE EDITOR

M. G. KENDALL

ISSUED BY

THE BIOMETRIKA OFFICE, UNIVERSITY COLLEGE, LONDON

PRINTED AT THE UNIVERSITY PRESS, CAMBRIDGE

[*Issued June* 1960]

This volume of *Biometrika* is published with the co-operation of

F. N. DAVID	N. L. JOHNSON
J. DURBIN	D. V. LINDLEY
M. J. R. HEALY	R. L. PLACKETT

A volume containing about 500 pages will be published annually in two half-yearly issues, appearing in June and December.

Papers for publication should be sent to

PROFESSOR E. S. PEARSON

Department of Statistics, University College London,
Gower Street, London, W.C. 1

It is a condition of publication in *Biometrika* that the paper shall not already have been issued elsewhere, and will not be reprinted without leave of the Editors.

Contributors receive 25 copies of their papers free. Joint authors 15 copies each. Order forms for separates are sent to authors with proofs of their papers.

SUBSCRIPTIONS for 1960 (Volume 47)

The subscription price, payable in advance is:
£2. 14*s.* (or **$8.00**) net per volume *including packing and postage*

Subscribers who buy the journal *Biometrika* for their own and not for general use may apply to the Secretary to be registered as **Personal Subscribers,** for whom the subscription price is **£2. 5*s.*** (payable only in sterling).

Cheques should be made payable to *Biometrika*, crossed 'a/c BIOMETRIKA TRUST' and sent to THE SECRETARY, BIOMETRIKA OFFICE, UNIVERSITY COLLEGE, GOWER STREET, LONDON, W.C. 1, to whom all orders for Offprints, Index and Table Separates (see pages (iv) to (vi)) should also be sent. For particulars as to the availability of **Back Issues** please see inside back cover.

[Members of the Institute of Mathematical Statistics are reminded that they may subscribe to *Biometrika* through the Institute who will accept subscriptions to *Biometrika* at the special rate of $6.50 if paid to the Treasurer of the Institute before 1 March.]

Biometrika (1960), **47**, 1 *and* 2, *p.* 1
Printed in Great Britain

A comparison of theoretical and empirical results for some stochastic population models*

17
173
8813A

BY M. S. BARTLETT, J. C. GOWER AND P. H. LESLIE

Statistical Laboratory, University of Manchester; Statistical Department, Rothamsted Experimental Station; Bureau of Animal Population, Department of Zoological Field Studies, Oxford

1. GENERAL REMARKS

In recent papers Bartlett (1957), Leslie (1958) and Leslie & Gower (1958) have illustrated by means of artificial series the properties of various idealized models of biological systems, including the single-species logistic stochastic process and two-species extensions. While these artificial series are always useful in an auxiliary qualitative sense, the theoretical intractability of many of these models has given the artificial series a somewhat more important role than they might otherwise have had. Nevertheless, what theoretical results there are should not be neglected, and indeed some of these were compared with empirical results from series in the last two of the papers mentioned above. It is the purpose of the present paper to indicate somewhat more systematically where theoretical results, even when only approximate, may be useful, and to make some further comparisons with the empirical results available.

PART I. THEORETICAL RESULTS

2. SINGLE-SPECIES MODELS (CONTINUOUS TIME)

Consider first a stochastic population model for a single species, with transition probabilities (in continuous time during the infinitesimal interval dt) $\lambda_n dt$ of a 'birth', and $\mu_n dt$ of a 'death', where n is the total population size. A 'death' may include emigration, but unless a 'birth' can include immigration, $\lambda_0 = 0$. If $\lambda_0 = 0$, an ultimate stationary distribution for n cannot strictly exist, but may effectively exist over all realizable time-intervals (see § 5; also Leslie (1958), Bartlett (1960)). Under conditions for which a stationary (or quasi-stationary) distribution does exist, the probability distribution for it must satisfy the recurrence relation

$$\mu_n P(n) = \lambda_{n-1} P(n-1) \tag{1}$$

(see, for example, Bartlett, 1960), from which relation the exact distribution $P(n)$ may always be calculated numerically, as will be illustrated below (§ 6). Under some further conditions which include m, $m/\sigma \gg 1$, we have asymptotically

$$P(n) \sim C \exp\{-\tfrac{1}{2}(n-m)^2/\sigma^2\}, \tag{2}$$

where m is the relevant solution of $\lambda_m = \mu_m$, and

$$\sigma^2 = -1 \Big/ \left[\frac{d(\lambda_n/\mu_n)}{dn}\right]_{n=m}. \tag{3}$$

* The work of one of the authors (M.S.B.) was supported in part by a research contract between the Office of Naval Research and the Department of Statistics, Harvard University.

If alternatively the properties of small fluctuations about the mean of the stationary distribution are investigated directly, the results obtained to the first approximation will be equivalent to the normal approximation above. As the first mention of this approach (Bartlett, 1956) was rather brief, it seems worth while showing how it may be developed to include second-stage (or even higher order) corrections. The procedure is sufficiently illustrated by means of the logistic model

$$\lambda_n = a_1 n - b_1 n^2, \quad \mu_n = a_2 n + b_2 n^2, \tag{4}$$

where λ_n remains zero when $n > a_1/b_1$, $(a_1, b_1, a_2, b_2 > 0)$. The stochastic equation (cf. Bartlett, 1957) is

$$dN_t = (\lambda_N - \mu_N)\,dt + dZ_1 - dZ_2, \tag{5}$$

or

$$N_{t+dt} = N_t + (\lambda_N - \mu_N)\,dt + dZ_1 - dZ_2. \tag{6}$$

Averaging (6) on the assumption that a stationary distribution has been reached, we have (from the coefficient of dt)

$$(a_1 - a_2)\,m - (b_1 + b_2)\,(\sigma^2 + m^2) = 0, \tag{7}$$

where $m \equiv E\{N\}$, $\sigma^2 \equiv E\{(N-m)^2\}$. Write further $\delta N_t = N_t - m$; then

$$\delta N_{t+dt} = \delta N_t + (\lambda_N - \mu_N)\,dt + dZ_1 - dZ_2. \tag{8}$$

Squaring and averaging this equation, we have exactly

$$2[(a_1 - a_2)\,\sigma^2 - (b_1 + b_2)\,\mu_3] + [(a_1 + a_2)\,m - (b_1 - b_2)\,(m^2 + \sigma^2)] = 0,$$

where $\mu_3 \equiv E\{(\delta N)^3\}$. Hence to the first order of approximation (noting that

$$m \sim (a_1 - a_2)/(b_1 + b_2), \quad \mu_3 \sim 0$$

to this order)

$$\sigma^2 \sim (a_1 - b_1 m)/(b_1 + b_2). \tag{9}$$

Similarly from the cube of (8) we obtain, noting that the averaged cube of $dZ_1 - dZ_2$ is strictly zero at $N = m$,

$$3E\{(\lambda_N - \mu_N)\,dt(\delta N)^2\} + 3E\{(dZ_1 - dZ_2)^2\,\delta N\} = 0.$$

From the normality approximation, we can to the second order of approximation write in this equation $\mu_4 \equiv E\{(\delta N)^4\} = 3\sigma^4$. We thus obtain

$$(a_1 - a_2)\,(\mu_3 + m\sigma^2) - (b_1 + b_2)\,(m^2\sigma^2 + 2m\mu_3 + 3\sigma^4) = -(a_1 + a_2)\,\sigma^2 + (b_1 - b_2)\,(2m\sigma^2 + \mu_3),$$

whence to the same order, as $m \gg 1$,

$$\mu_3 m(b_1 + b_2) \sim \sigma^2 m(b_2 - b_1),$$

that is

$$\mu_3 \sim \sigma^2(b_2 - b_1)/(b_2 + b_1). \tag{10}$$

It will be seen that μ_3 is only zero to this order if $b_1 = b_2$; and it changes sign as we move from a constant birth-rate ($b_1 = 0$) to a constant death-rate ($b_2 = 0$).

3. SINGLE-SPECIES MODELS (DISCRETE TIME)

Before, however, we consider comparing any of these results with empirical results obtained from Leslie's artificial series, we must recall that the latter were obtained on the basis of a discrete-time model (Leslie, 1958), which has its own theoretical distribution. Whilst its exact form would be complicated, the investigation of approximative moment

formulae proceeds very similarly, as was noted by Leslie & Gower (1958) in the case of a two-species model. We shall illustrate the procedure here for a single-species model, taking it to the next stage of approximation.

The transitions in Leslie's model are obtained from the recurrence formulae

$$\left.\begin{aligned} E\{N_{t+1}\} &= e^{(b_t-d_t)}\,N_t, \\ \sigma^2\{N_{t+1}\} &= \frac{b_t+d_t}{b_t-d_t}(e^{b_t-d_t}-1)\,e^{b_t-d_t}\,N_t, \end{aligned}\right\} \tag{11}$$

treating b_t and d_t constant from t to $t+1$, and assuming also ΔN_t normal (with the restriction $N_{t+1} > 0$). Putting $b_t - d_t = \log \lambda_t$, we have in the recurrence relation

$$N_{t+1} = f(N_t) + Z_{t+1}, \tag{12}$$

$f(N_t) = \lambda_t N_t$ in this case, where λ_t for the logistic model is of the form $\lambda/(1+\alpha N_t)$. Putting $\delta N_t = N_t - m$, where $m = E\{N_t\}$ under stationary conditions, and writing

$$f(N_t) = f(m) + \delta N_t \frac{\partial f}{\partial m} + \tfrac{1}{2}(\delta N_t)^2 \frac{\partial^2 f}{\partial m^2} + \ldots,$$

where $\partial f/\partial m$ denotes $\partial f/\partial N_t$ at the value $N_t = m$, etc., we have in the first approximation

$$m = f(m) = (\lambda - 1)/\alpha, \tag{13}$$

$$\left.\begin{aligned} \sigma^2 &= \frac{\sigma_Z^2}{1-(\partial f/\partial m)^2} = \frac{\sigma_Z^2}{1-(1/\lambda)^2}, \\ \mu_3 &= 0. \end{aligned}\right\} \tag{14}$$

To the next approximation $\quad f(m) - m + \dfrac{1}{2}\dfrac{\partial^2 f}{\partial m^2}\sigma^2 = 0,$

where $\partial f/\partial m = 1/\lambda$, $\partial^2 f/\partial m^2 = -2\alpha/\lambda^2$; and

$$\mu_3 = E\left\{\left(\delta N_t \frac{\partial f}{\partial m} + \tfrac{1}{2}[(\delta N_t)^2 - \sigma^2]\frac{\partial^2 f}{\partial m^2} + Z_{t+1}\right)^3\right\} \sim \mu_3(Z) + \left(\frac{\partial f}{\partial m}\right)^3 \mu_3$$
$$+ \frac{3}{2}\left(\frac{\partial f}{\partial m}\right)^2 \frac{\partial^2 f}{\partial m^2}(\mu_4 - \sigma^4) + 3E\left\{\sigma^2(Z_{t+1} \mid N_t)\left(\delta N_t \frac{\partial f}{\partial m} + \tfrac{1}{2}[(\delta N_t)^2 - \sigma^2]\frac{\partial^2 f}{\partial m^2}\right)\right\}.$$

Now $\mu_3(Z) = 0$, $\mu_4 \sim 3\sigma^4$, and it only remains to evaluate such terms as $E\{\sigma^2(Z_{t+1} \mid N_t)\,\delta N_t\}$. It should be noted that the value of this expression may depend on the precise numerical procedure adopted in obtaining the artificial series. Thus if $\sigma^2(Z_{t+1} \mid N_t)$ were taken constant, say at the value $\sigma^2(Z_{t+1} \mid m)$, the whole of the last term in the expression above for μ_3 would be zero. However, if, more accurately, we expand $\sigma^2(Z_{t+1} \mid N_t)$ in (11) in the neighbourhood of $N_t = m$ (where $b_t \sim d_t$), we find

$$\sigma^2(Z_{t+1} \mid N_t) \sim (2b - (3b-1)\,\alpha\delta N_t/\lambda)\,N_t \tag{15}$$

when the birth-rate is constant (b), and

$$\sigma^2(Z_{t+1} \mid N_t) \sim (2d - (3d+1)\,\alpha\delta N_t/\lambda)\,N_t \tag{16}$$

when the death-rate is constant (d). These results, incidentally, may be useful as approximations for $\sigma^2(Z_{t+1})$ when artificial series are being constructed. Thus, retaining terms of the appropriate order, we find for $b_t \equiv b$,

$$\mu_3(1-(1/\lambda)^3) \sim \frac{3\sigma^2(\lambda-1)}{\lambda^2}\left\{\frac{(2b)\,(\lambda+2)}{\lambda+1} - (3b-1)\right\}. \tag{17}$$

Similarly in the case of constant death-rate (d),

$$\mu_3(1-(1/\lambda)^3) \sim \frac{3\sigma^2(\lambda-1)}{\lambda^2}\left\{\frac{(2d)(\lambda+2)}{\lambda+1}-(3d+1)\right\}. \tag{18}$$

It may be useful to summarize the formulae derived in §§ 2 and 3.

Logistic model; continuous time

$\lambda_n = a_1 n - b_1 n^2$; $\mu_n = a_2 n + b_2 n^2$.
Variance (1st approx.) $\sigma^2 \sim (a_1 - b_1 m)/(b_1 + b_2)$.
Mean (2nd approx.) $m' \sim m - \sigma^2/m$, where $m = (a_1 - a_2)/(b_1 + b_2)$.
Skewness (2nd approx.) $\mu_3 \sim \sigma^2(b_2 - b_1)/(b_2 + b_1)$.

Discrete time

$\lambda_t = \lambda/(1+\alpha N_t)$, (i) birth-rate constant, b; (ii) death-rate constant, d.

(i) Variance (1st approx.) $\sigma^2 \sim 2bm/\{1-(1/\lambda)^2\}$, where $m = (\lambda-1)/\alpha$.
Mean (2nd approx.) $m' \sim m - \sigma^2/(\lambda m)$.
Skewness (2nd approx.)

$$\mu_3 \sim \frac{3\sigma^2\lambda(\lambda-1)}{\lambda^3-1}\left\{\frac{2b(\lambda+2)}{\lambda+1}-(3b-1)\right\}.$$

(ii) Variance (1st approx.) $\sigma^2 \sim 2dm/\{1-(1/\lambda)^2\}$.
Mean (2nd approx.) $m' \sim m - \sigma^2/(\lambda m)$.
Skewness (2nd approx.)

$$\mu_3 \sim \frac{3\sigma^2\lambda(\lambda-1)}{\lambda^3-1}\left\{\frac{2d(\lambda+2)}{\lambda+1}-(3d+1)\right\}.$$

4. MODELS WITH TWO SPECIES

Whilst the above methods are available for models with two (or more) species, the formulae get rather complicated; as the first approximation results in the case of a quasi-stationary distribution for two species have already been indicated by Leslie & Gower (1958), they will not be listed here. With regard to the exact recurrence relation (1) for the distribution $P(n)$, the corresponding relation for two species is easy to write down, but has no simple method of solution. However, under conditions for which a well-defined stationary (or quasi-stationary) distribution exists with m_i, $m_i/\sigma_i \gg 1$, the distribution will be approximately bivariate normal, with its moments given by the approximate formulae already referred to.

In the case of models for which one species will become extinct, it is also easy to write down the equation satisfied by the extinction probability, say, $p(n, n')$ for the first species, if n and n' are the initial numbers of the two species. Thus consider a continuous time model with birth- and death-rates:

1st species	2nd species
$\lambda(n,n') = a_1 - b_1 n - c_1 n'$	$\lambda'(n,n') = a_1' - b_1' n' - c_1' n$
$\mu(n,n') = a_2 + b_2 n + c_2 n'$	$\mu'(n,n') = a_2' + b_2' n' + c_2' n$

The equation for $p(n, n')$ is obtained from the differential equation for the chance of extinction $p_t(n, n')$ by time t, and letting $t \to \infty$; this equation for $p_t(n, n')$ is readily derived from the possible transitions in the *first* infinitesimal interval dt. We obtain

$$n\lambda(n,n')[p(n+1,n')-p(n,n')]+n\mu(n,n')[p(n-1,n')-p(n,n')]$$
$$+n'\lambda'(n,n')[p(n,n'+1)-p(n,n')]+n'\mu'(n,n')[p(n,n'-1)-p(n,n')]=0, \quad (19)$$

with boundary conditions $p(n,0) = 0, p(0,n') = 1$. While it is possible to solve this equation, for example, by iteration, in any actual example, it seems quicker to obtain approximate answers by Monte Carlo methods, as shown by Leslie & Gower (1958).

5. Recurrence times

If it is desired to check recurrence times to any state S in a stationary process with a finite number of states, the relevant formulae have been given by Bartlett (1955, § 6.41). Thus in the case of discrete time, the mean recurrence time is

$$\Theta_1 = \frac{1-P(S)}{P(S)[1-P(S \mid S)]}, \quad (20)$$

where $P(S \mid S)$ denotes the conditional probability of S at one instant, given S at the previous instant. In the case of continuous time, if $P(S \mid S)$ for times separated by an interval δt is $1-\epsilon\,\delta t+o(\delta t)$, then

$$\Theta_1 = \frac{1-P(S)}{\epsilon P(S)}. \quad (21)$$

This formulae is relevant in assessing the passage-time to the zero state S for population processes with an 'absorbing barrier' at this state, for if we insert a fictitious escape probability $\epsilon\,\delta t$ in δt, we have from the equilibrium for the state $N = 1$,

$$\mu_1 P(1) = \epsilon P(0), \quad (22)$$

whence

$$\Theta_1 = \frac{1-P(0)}{\mu_1 P(1)}. \quad (23)$$

Under conditions for which $P(0)$ is small, the quasi-stationary distribution $P(n)$, $(n > 0)$, exists approximately independently of ϵ, and under such conditions the mean recurrence time $\Theta_1 \sim 1/[\mu_1 P(1)]$ gives the order of magnitude of the passage-time to zero. For processes of the type discussed in § 2 such passage-times may be so large as to be considered infinite (cf. Leslie, 1958).

PART II. NUMERICAL RESULTS

6. Example of distribution $P(n)$

It was noted in § 2 that the recurrence relation for $P(n)$ enabled $P(n)$ to be calculated exactly. This is strictly true only if the zero state is not absorbing, but from the last section $P(n)$ is effectively defined in quasi-stationary cases also. Whilst it is still a purely theoretical result, we give a numerical example of $P(n)$ in this last case, for the logistic model

$$a_1 = 0{\cdot}8077, \quad b_1 = 0{\cdot}006932,$$
$$a_2 = 0{\cdot}1145, \quad b_2 = 0.$$

$P(n)$, standardized to a total of 4975, was found to have the values given in Table 1.

Table 1 (giving $f = 4975\,P(n)$)

n	f	n	f	n	f
81	0·1	93	118·3	105	236·7
82	0·2	94	166·7	106	163·4
83	0·5	95	224·7	107	103·1
84	1·0	96	289·7	108	58·8
85	1·9	97	356·0	109	30·1
86	3·7	98	416·3	110	13·5
87	6·7	99	461·8	111	5·3
88	11·8	100	484·8	112	1·7
89	20·2	101	480·1	113	0·5
90	33·2	102	446·8	114	0·1
91	52·7	103	388·9		
92	80·6	104	315·1	Total	4975·0

The constants of this distribution as calculated by the approximate formulae of § 2 agree very well with the exact results:

	Exact	Approximation
m:	99·83	99·83 (2nd approx.)
σ^2:	16·71	16·52 (1st approx.)
μ_3:	−16·61	−16·52 (2nd approx.)

7. Comparison of approximate moments for discrete-time model with empirical results

Four empirical distributions for a logistic process in the region of the stationary state were built up on the Elliott—N.R.D.C. 401 computer at Rothamsted Experimental Station, taking the values of the parameters in the discrete-time model

$$E\{N_{t+1} \mid N_t\} = \frac{\lambda}{1+\alpha N_t} N_t = \lambda_t N_t \quad (\lambda = e^{b-d}),$$

given in Table 2. In each case $m = (\lambda - 1)/\alpha = 100$, and it will be noted that the unit of time in Ib is $\frac{1}{5}$th of that adopted in the remaining three models.

Table 2

Model	λ	α	Constant birth-rate (b)	Constant death-rate (d)
Ia	2·0	0·01	1·0083	—
Ib	1·1487	0·001487	0·2017	—
IIa	2·0	0·01	—	0·1145
IIb	2·0	0·01	—	0·3151

The programmes used for computing Ia, IIa and IIb were originally written for a system of two competing species (Leslie & Gower, 1958), but by putting two of the parameters equal to zero (cf. the equations given later in § 8), these could be used for computing simultaneously a pair of logistic processes. In order to simplify these programmes, however, certain approximations had been made to var $(N_{t+1} \mid N_t)$. Thus, if the expression for the variance in (11) is written as

$$\sigma^2\{N_{t+1}\} = \phi E\{N_{t+1}\},$$

where, when the birth-rate (b) remains constant (BRC model),

$$\phi = \left(\frac{2b}{r_t} - 1\right)(\lambda_t - 1), \tag{24}$$

and when the death-rate (d) remains constant (DRC model),

$$\phi = \left(\frac{2d}{r_t}+1\right)(\lambda_t-1), \tag{25}$$

then it may be shown empirically (Leslie, 1958), by tabulating ϕ over a relatively wide range of possible values of λ_t (or $\log_e \lambda_t = r_t$), that in the BRC model we have, when $\lambda = 2{\cdot}0$ and constant $(b) = 1{\cdot}0083$,

$$\sigma^2\{N_{t+1}\} \sim 2E\{N_{t+1}\} \tag{26}$$

and in the DRC model for

$$\left.\begin{aligned} &\text{constant } (d) = 0{\cdot}1145: \quad \sigma^2\{N_{t+1}\} \sim (-0{\cdot}87+1{\cdot}10\lambda_t)\,E\{N_{t+1}\},\\ &\text{constant } (d) = 0{\cdot}3151: \quad \sigma^2\{N_{t+1}\} \sim (-0{\cdot}66+1{\cdot}29\lambda_t)\,E\{N_{t+1}\}.\end{aligned}\right\} \tag{27}$$

These empirical approximations, which should hold over the entire development of any process with the given parameters, are closely related to (15) and (16) for systems in the neighbourhood of the stationary mean. For, expressing the latter in terms of $E\{N_{t+1} \mid N_t\}$, we have in the region of $N_t \sim m$,

$$N_t \sim \left(1+\frac{\alpha}{\lambda}\delta N_t\right) E\{N_{t+1}\}, \quad \lambda_t \sim 1-\frac{\alpha}{\lambda}\delta N_t,$$

and (15) and (16) can be written, respectively, as

$$\sigma^2\{Z_{t+1} \mid N_t\} \sim [(b+1)+(b-1)\lambda_t]\,E\{N_{t+1}\}, \tag{28}$$

$$\sigma^2\{Z_{t+1} \mid N_t\} \sim [(d-1)+(d+1)\lambda_t]\,E\{N_{t+1}\}. \tag{29}$$

Thus, in the BRC model, when $b \sim 1$,

$$\sigma^2\{Z_{t+1} \mid N_t\} \sim 2E\{N_{t+1}\},$$

while in the DRC model, for $d = 0{\cdot}1145$,

$$\sigma^2\{Z_{t+1} \mid N_t\} \sim [-0{\cdot}8855+1{\cdot}1145\lambda_t]\,E\{N_{t+1}\},$$

and for $d = 0{\cdot}3151$

$$\sigma^2\{Z_{t+1} \mid N_t\} \sim [-0{\cdot}6849+1{\cdot}3151\lambda_t]\,E\{N_{t+1}\},$$

which correspond very closely to (27).

In the remaining model (Ib), a programme was used in which the 'exact' expression (24) for ϕ was incorporated. This has the advantage that the choice of the constant birth-rate (b) is not restricted to only a very limited range of values, as in the case of the approximation (26).

The observed moments of these computed distributions, together with those expected from the theoretical approximations for the discrete-time (D-T) and equivalent continuous-time (C-T) models, are shown in Table 3.

The agreement between the observed moments and the theoretical approximations for the discrete-time model seems very satisfactory, considering the errors involved in estimating the former, even in samples of this size. The standard errors quoted are in general classical values, ignoring the serial correlations between the successive observations, and representing lower limits to the correct values. The correlation ρ_1 between successive observations is to the first approximation $1/\lambda$, and the correcting factors to the standard errors of the mean and standard deviation are respectively $\sqrt{\{(1+\rho_1)/(1-\rho_1)\}}$ and $\sqrt{\{(1+\rho_1^2)/(1-\rho_1^2)\}}$. Where the differences between the observed and theoretical means or standard deviations

exceed twice the classical standard error, the corrected standard error is shown in brackets, and it will be seen that the one such difference no longer appears anomalous. The corrections to the standard errors of moments to allow for dependence become more complicated for the higher moments, but are not needed for the γ_1 values in Table 3 (formulae for them may be ascertained if required from Chanda (1958)).

Table 3

Model	No. of observations	Moment	Approximations (C-T) model	Approximations (D-T) model	Observed (D-T) model
I*a*	9950	Mean	98·54	98·66	98·60 ± 0·16
		σ	12·06	16·40	16·47 ± 0·12
		γ_1	0·083	0·035	0·072 ± 0·025
I*b*	5240	Mean	98·54	98·55	98·63 ± 0·17
		σ	12·06	12·90	12·53 ± 0·12 (0·33)
		γ_1	0·083	0·076	0·053 ± 0·034
II*a*	4975	Mean	99·83	99·85	99·97 ± 0·08
		σ	4·06	5·53	5·51 ± 0·06
		γ_1	−0·246	−0·161	−0·182 ± 0·035
II*b*	4975	Mean	99·55	99·58	99·44 ± 0·13
		σ	6·74	9·17	9·16 ± 0·09
		γ_1	−0·148	−0·103	−0·103 ± 0·035

Table 4

N	$f(N)$	N	$f(N)$	N	$f(N)$
77	1	92	138	107	188
78	0	93	166	108	120
79	0	94	194	109	90
80	1	95	213	110	64
81	2	96	270	111	54
82	5	97	285	112	25
83	3	98	310	113	17
84	4	99	329	114	14
85	9	100	350	115	6
86	20	101	361	116	0
87	33	102	357	117	0
88	35	103	339	118	1
89	51	104	318		
90	67	105	261		
91	97	106	177	Total	4975

As an illustration, we give in Table 4 the observed frequency distribution for model II*a*, corresponding to the continuous-time model for which the exact distribution is given in § 6.

The main difference between this observed distribution and the exact form for the equivalent continuous-time model is in the scale of the variance σ^2 (one notes in passing that in both cases $\mu_3 \sim -\sigma^2$). It is, however, always possible to make the variances of the two types of model more in agreement by adopting a smaller unit of time in the discrete-time model, as is illustrated in the cases of I*a* and I*b* in the above table.

It may also be of interest to consider the recurrence times which were observed in this set of realizations for model II*a*. Regarding the occurrence of a particular integer as a specified state S being occupied, the mean life-time spent in the state S, $T_1 = 1/\{1 - P(S \mid S)\}$,

could be determined from the typed lists of results, and hence the mean recurrence time for the state S,

$$\Theta_1 = T_1 \frac{1-P(S)}{P(S)}.$$

The observed values of T_1 and Θ_1, neglecting the tails of the distribution where the observed frequencies became small, are given in Table 5.

Table 5

State S	Observed T_1	Observed Θ_1	Normal approximation Θ_1	State S	Observed T_1	Observed Θ_1	Normal approximation Θ_1
85	1·1250	623·0	551·7	100	1·0870	14·3	13·8
86	1·0000	247·7	343·2	101	1·0841	13·9	14·1
87	1·0312	154·5	220·5	102	1·0593	13·7	14·8
88	1·0294	145·2	146·4	103	1·1042	15·1	16·1
89	1·0851	104·7	100·5	104	1·1042	16·1	18·0
90	1·0635	78·0	71·2	105	1·0830	19·6	20·9
91	1·0211	51·4	52·2	106	1·0727	29·1	25·0
92	1·0455	36·7	39·5	107	1·1124	28·4	30·9
93	1·0573	30·7	30·9	108	1·0526	42·6	39·5
94	1·0838	26·7	25·0	109	1·0465	56·8	52·2
95	1·1036	24·7	20·9	110	1·0159	77·9	71·2
96	1·0887	18·9	18·0	111	1·0000	91·1	100·5
97	1·0634	17·5	16·1	112	1·0870	215·2	146·4
98	1·0954	16·4	14·8	113	1·0625	309·8	220·5
99	1·0717	15·1	14·1	114	1·0000	354·8	343·2

If the conditional probability of S, given S at the preceding instant, $P(S \mid S) \sim P(S)$, as appears very roughly to be the case in this example (e.g. for the ten states disposed symmetrically about the mean, the average $P(S \mid S) = 0{\cdot}0785$ and the average $P(S) = 0{\cdot}0630$), then

$$\Theta_1 \sim 1/P(S).$$

Taking the first approximations to the moments of the discrete-time distribution, when the death-rate remains constant, viz.

$$\text{mean} = (\lambda - 1)/\alpha = m,$$
$$\sigma^2 = 2dm/\{1-(1/\lambda)^2\},$$
$$\mu_3 = 0,$$

then the normal approximation is

$$\Theta_1 \sim \sqrt{(2\pi)}\,\sigma \exp\left[\frac{(N-m)^2}{2\sigma^2}\right].$$

These figures are given in the third column of the above table, and it will be seen that although the departures from normality of the actual distribution are appreciable, yet on the whole the approximation indicates the order of magnitude of the observed recurrence times, more particularly for the states which are less than the mean value. By extrapolation

$$1/P(1) \sim 7{\cdot}0 \times 10^{70},$$

and it is evident, without proceeding any further, and without taking the approximate value too literally, that the probability of random extinction for this system is negligible, even in the case of the discrete-time model with its larger variance.

8. COMPARISON OF APPROXIMATE MOMENTS WITH EMPIRICAL RESULTS FOR A TWO-SPECIES SYSTEM

Two bivariate distributions were computed for a system of two competing species fluctuating in the region of the stable stationary state. It was assumed in both cases that the death-rate (d) of each species remained constant, and in the set of deterministic equations defining the expectations for this type of system (Leslie & Gower, 1958),

$$E\{N_1(t+1) \mid N_1(t), N_2(t)\} = \frac{\lambda_1 N_1(t)}{1+\alpha_1 N_1(t)+\beta_1 N_2(t)}, \tag{30}$$

$$E\{N_2(t+1) \mid N_1(t), N_2(t)\} = \frac{\lambda_2 N_2(t)}{1+\alpha_2 N_2(t)+\beta_2 N_1(t)}, \tag{31}$$

the following parameters were adopted in System A,

$$\lambda_1 = 2{\cdot}5, \quad d_1 = 0{\cdot}1145, \quad \alpha_1 = 0{\cdot}008, \quad \beta_1 = 0{\cdot}003,$$
$$\lambda_2 = 2{\cdot}0, \quad d_2 = 0{\cdot}3151, \quad \alpha_2 = 0{\cdot}00625, \quad \beta_2 = 0{\cdot}0025.$$

In the second case, System B, the same set of parameters was used, except that the value of α_1, was changed to $\alpha_1 = 0{\cdot}005$. The computed distributions were based on 1000 observations for A, and on 995 for B.

The first approximations to the moments are (assuming that both N_1 and N_2 are distributed normally about the stationary state)

$$\text{mean}\,(N_1) = \frac{\alpha_2(\lambda_1-1)-\beta_1(\lambda_2-1)}{\alpha_1\alpha_2-\beta_1\beta_2}, \tag{32}$$

$$\text{mean}\,(N_2) = \frac{\alpha_1(\lambda_2-1)-\beta_2(\lambda_1-1)}{\alpha_1\alpha_2-\beta_1\beta_2}, \tag{33}$$

and for the marginal distributions,

$$\mu_3(N_1) = \mu_3(N_2) = 0; \tag{34}$$

while the variances and covariances may be obtained from the solution of the equations which have already been given for the discrete-time model (Leslie & Gower, 1958, § 5). The results are shown in Table 6.

Table 6

	System A		System B	
	Approximation	Observed	Approximation	Observed
Mean (N_1)	150·0	150·82	268·4	270·16
Mean (N_2)	100·0	99·19	52·6	50·75
$\sigma(N_1)$	7·84	7·85	11·06	11·37
$\sigma(N_2)$	11·49	12·03	11·68	12·42
$\rho(N_1, N_2)$	−0·364	−0·446	−0·529	−0·602
$\gamma_1(N_1)$	0	+0·001	0	−0·005
$\gamma_1(N_2)$	0	+0·074	0	−0·022

The agreement between the approximations and the observed means, standard deviations and correlation coefficients seems very satisfactory in both cases; while the magnitude of the skewness coefficients suggest that both distributions could be regarded as approximately bivariate normal in form.

References

Bartlett, M. S. (1955). *An Introduction to Stochastic Processes.* Cambridge University Press.

Bartlett, M. S. (1956). Deterministic and stochastic models for recurrent epidemics. *Proc. Third Berkeley Symposium on Mathematical Statistics and Probability*, **4**, 81. University of California Press.

Bartlett, M. S. (1957). On theoretical models for competitive and predatory biological systems. *Biometrika*, **44**, 27–42.

Bartlett, M. S. (1960). Some stochastic models in ecology and epidemiology. Published in the dedicatory volume to Professor Harold Hotelling.

Chanda, K. C. (1958). Sampling properties of tests for goodness-of-fit, normality and stationarity of time series. Ph.D. Thesis, University of Manchester.

Leslie, P. H. (1958). A stochastic model for studying the properties of certain biological systems by numerical methods. *Biometrika*, **45**, 16–31.

Leslie, P. H. & Gower, J. C. (1958). The properties of a stochastic model for two competing species. *Biometrika*, **45**, 316–30.

Biometrika (1960), **47**, 1 *and* 2, *p.* 13
Printed in Great Britain

Birth-and-death processes, and the theory of carcinogenesis

BY DAVID G. KENDALL*
Magdalen College, Oxford

1. INTRODUCTION

In March 1959, Professor J. Neyman pointed out to me that some calculations of mine (Kendall, 1952) relating to phenotypically delayed mutations could be adapted so as to form a very rough quantitative model for a current theory of carcinogenesis, and he kindly suggested that I should explore this possibility and report on it to the San Diego session of the Biometric Society in June 1959. I wish to begin by expressing my gratitude to him for suggesting this problem and for sketching the lines along which it might be tackled.

We first outline the biological situation and describe the model which is intended to represent it. To begin with we suppose that we have a large (say infinite) population of normal cells subject to carcinogenic action (perhaps by irradiation) usually of two kinds; a 'background', which is always present, and an experimental enhancement of this of known intensity and acting over a known period. We suppose that the joint effect of these agents is to transform some normal cells into first-order mutants, which it will be convenient to call *grey* cells (simply in order to have a short name for them). Quantitatively we assume that in a time interval (t_1, t_2) the number of grey cells produced in this way is a Poisson variable of expectation

$$\int_{t_1}^{t_2} f(t)\,dt, \tag{1·1}$$

where $f(t)$ represents the intensity of the carcinogenic action (of both kinds). In most experimental situations $f(t)$ will be a step-function.

Each grey cell will be supposed to generate a clone (identified with a benign growth) the individuals in which multiply independently with birth rate λ and death rate μ as in the simple birth-and-death process (for which see, for example, Kendall, 1952). An important assumption is that μ is greater than λ; this ensures that the grey clones each become extinct, with probability one, if given sufficient time in which to do so. Thus *benign* growths are here being represented by *sub*critical birth-and-death processes.

It is next supposed that as a result of the same carcinogenic influences the grey cells are capable of being transformed into second-order mutants, which for ease of reference we shall call *black* cells. Two alternative assumptions can be made here:

(A) the effect of a second-order mutation is to transform a single grey cell into a black cell; such mutations occur independently of the birth-and-death process controlling the growth of the clone to which the grey cell belongs;

(B) second-order mutations occur (if at all) at epochs of cell division, and their effect is to convert *one* of the two fission products into a black cell, so that we have

$$\text{grey} \rightarrow \text{grey} + \text{black}.$$

We could (but will not now) add a third variation of the model, in which when a second-order mutation acts at the moment of a cell division, *both* fission products are black.

* This paper was prepared with the partial support of a research grant from the National Institutes of Health, U.S. Public Health Service.

Once a black cell has been formed, it generates a black clone developing according to a birth-and-death process, *but now the birth rate L and death rate M satisfy the reversed inequality*: $M < L$. This means that we are identifying *malignant* growths with *super*critical birth-and-death processes. The chance that a black clone eventually becomes extinct will be M/L; we could make this zero, if we wish, by setting $M = 0$. It is known, in any case, that if a population controlled by a supercritical birth-and-death process does not become extinct then its size grows without limit, ultimately exponentially.*

We shall assume that for a grey cell extant at the epoch t there is a chance $\lambda\,dt$ that it will split into two grey cells during $(t, t+dt)$ and that there is a chance $\mu\,dt$ that it will 'die' (and then cease to be numbered among the grey cells) during the same time period. (There is plainly an alternative assumption possible here; we could assume that after 'death' a grey cell is still available to be counted as grey by the experimenter, and that it is dead merely in that it no longer multiplies and is no longer available for conversion into a black cell. It would be quite easy to incorporate this assumption in the model, if desired. We shall indicate later how it could be done.)

We have still to specify the probability with which a second-order transition involving a given grey cell will take place in a time interval $(t, t+dt)$. If we are using Assumption (A), then we write $\nu\,dt$ for the chance that the grey cell will blacken in the stated interval. If we are using Assumption (B), we write $\nu\,dt$ for the chance that the grey cell will be converted into a grey cell plus a black cell during the stated interval. These probabilities are of course only intended to be correct up to terms in $o(dt)$, and multiple events are supposed to have probability $o(dt)$, so that in very small intervals the four possibilities are (i) a λ-transition, (ii) a μ-transition, (iii) a ν-transition, and (iv) no change.

The birth-and-death rates λ, μ, L, and M are supposed to be constants. We may or may not wish ν to be treated as constant. If it is time-dependent, the natural procedure would be to put $\nu = \nu_0 f(t)$, where f is the function already used to measure the intensity of the carcinogenic action at various times. Possibly the occurrence of second-order mutations will depend on the carcinogenic agents in a different way. For example, if the background agents can produce second-order mutants but the experimental agents cannot, then it would be more correct to treat ν as constant. All we can do for the moment is to keep in mind the possibility that ν may have to be time-dependent.

We have now explained the model of the *phenomenon*, but we still have to complete the model of the *experiment*. Presumably this will consist in the counting of black and grey clones after the end of the exposure period, but certainly not all clones will be counted. Some will be missed; it is more likely that small clones will be missed than that large ones will be; thirdly, we cannot suppose that the grey and black clones of equal size will be detected with equal frequency. We must therefore introduce a function $C_g(n)$ giving the probability that a grey clone of size n will be counted by the experimenter, and a similarly defined function $C_b(n)$ for the black clones. Neyman has suggested that one might take

$$C_g(n) = 1-\gamma^n, \quad C_b(n) = 1-\beta^n, \tag{1·2}$$

where β and γ are positive numbers less than unity. This is a plausible assumption, because we might argue that if β is the chance that a single black cell will be missed, then β^n will be the chance that a clone of size n will be missed, and $1-\beta^n$ will be the chance that it will be

* This is most readily established (as I think T. E. Harris has pointed out) by observing that the population-size divided by the exponential of $(L-M)\,t$ yields a positive martingale.

counted—*if* (an unrealistic assumption) we grant that the individuals in a clone can be treated independently in this way. Bearing in mind that the experimenter will see not a complete clone but a two-dimensional section of one, it might seem better to replace n by the two-thirds power of n in formulae (1·2). Neyman reports that the formulae (1·2) do not fit very well, being not steep enough, but they have one immense theoretical advantage: they fit in admirably with the (inevitable) employment of generating functions. The two-thirds power law does not have this convenient property and will not be further considered here. In fact we shall work exclusively with the Neyman laws (1·2), so as to make the maximum of analytic progress. At a later stage this is a feature of the model which will probably require modification.

Another modification which may be needed is the abandonment of the simple negative-exponential law,

$$e^{-\tau}\,d\tau \quad (0 < \tau < \infty), \tag{1·3}$$

for the time intervals between the successive events (of course there is always a parameter to be inserted in (1·3) so as to give the correct *mean* time interval). It may be much better to use multiple-phase birth-and-death processes of the type investigated by Waugh (1955) and by myself (1952) to represent the growth of the grey and black clones. This, however, would greatly complicate the mathematics, and for the present purposes the model described above is the natural one to study. It is the simplest of the various models proposed by Neyman (1958) in his original formulation of the problem.*

2. The case of a single grey clone

To start with we shall consider the history of a single grey clone descended from a single grey cell formed at the epoch $t = 0$. At a subsequent epoch t, we will have say g_t grey *cells* and B_t black *clones* (of various sizes). Consider one of these black clones, formed say at the epoch u (where $0 < u < t$); it will have been growing for a period $(t-u)$ and if its size is b then we shall have (formula (81) of Kendall, 1952)

$$E\{z^b\} = \frac{M(T-1)-(MT-L)z}{(LT-M)-L(T-1)z} \quad \text{where} \quad T = \exp[(L-M)(t-u)]. \tag{2·1}$$

The chance that it will be counted, given that its size is b, is $1-\beta^b$, and so (averaging over the b-distribution) the final chance that it will be counted is

$$q(t-u) = \frac{(L-M)(1-\beta)T}{(LT-M)-L(T-1)\beta}, \quad \text{where} \quad T = \exp[(L-M)(t-u)]. \tag{2·2}$$

When $(t-u)$ (the life-time of the black clone) tends to infinity then this chance tends to the limit $q(\infty) = 1 - M/L$ (which is just the chance that the black clone will become infinitely large instead of becoming extinct).

We now introduce the double generating function

$$\phi(z, w, t) = E\{z^{g_t} w^{B'_t}\} \tag{2·3}$$

for the joint distribution of the number of grey cells (g_t) and the number of *detected* black clones (B'_t). We easily obtain the following integral equation for the function ϕ, by considering (a) whether or no an event occurs in $(0, t)$, (b) the epoch u of the first event (if there is one), (c) the nature of the first event, and (d) what happens after the first event. *We are here*

* For other stochastic models of tumour-formation see Arley & Iverson (1952) and also Armitage & Doll (1957).

using Assumption (A). We shall indicate the effect of employing Assumption (B) in a moment. If ϕ_{t-u} denotes $\phi(z, w, t-u)$, then the integral equation can be written*

$$\phi(z,w,t) = z\exp[-(\lambda+\mu+\nu)t] + \int_0^t \{\lambda\phi_{t-u}^2 + \mu + \nu[1-(1-w)q(t-u)]\} \times \exp[-(\lambda+\mu+\nu)u]\,du. \quad (2{\cdot}4\,\mathrm{A})$$

In writing out the integral equation *we have also assumed that ν is constant.* If it were not constant, then νt in the first and νu in the second exponential would have to be replaced by

$$\int_0^t \nu(\tau)\,d\tau, \quad \int_0^u \nu(\tau)\,d\tau,$$

respectively, and ν within the braces would have to be replaced by $\nu(u)$. This would be a tiresome complication, and to keep things as simple as possible we shall throughout the rest of this paper suppose ν constant. (This means that either (i) we suppose that only the background can produce second-order mutations or (ii) we keep the experimental carcinogenic agent at a constant intensity throughout the experiment.)

We now transform (2·4 A)† by changing the time variable from u to $\tau = t-u$, by multiplying through by the exponential of $(\lambda+\mu+\nu)t$ and by carrying out a differentiation with regard to t. We obtain

$$D\phi = \lambda\phi^2 - (\lambda+\mu+\nu)\phi + \mu + \nu[1-(1-w)q(t)], \quad (2{\cdot}5\,\mathrm{A})$$

with the side-condition

$$\phi(z,w,0) = z; \quad (2{\cdot}6\,\mathrm{A})$$

here D denotes d/dt.

This is a differential equation of the Riccati type for ϕ, but unfortunately it does not have constant coefficients. If we had used Assumption (B) instead of (A) then the term $\nu[1-(1-w)q(t)]$ would have had an extra factor ϕ, so that we should still have had a Riccati equation. In what follows we shall continue to assume both (A) and the constancy of ν.

Three things can be done with the equation (2·5 A) (apart from solving it—or the equivalent (2·4 A)—numerically). First of all we can easily find the joint moments of g_t and B_t' by differentiating (2·5 A) as many times as may be appropriate with regard to z and w, and then putting $z = w = 1$. The term ϕ^2 and its successors in this procedure will all be linearized, and one obtains a sequence or rather a double sequence of *linear* differential equations for product moments of various orders for these two random variables. The process is quite elementary, but we will not carry it out here because we shall make no use of the results.

Next, we can consider an approximation to (2·5 A) in which we replace $q(t)$ by its limit $q(\infty) = 1 - M/L$ (approached when $t \to \infty$). We then have

$$D\phi = \lambda\phi^2 - (\lambda+\mu+\nu)\phi + (\mu+\nu\bar{q}), \quad \text{where} \quad \bar{q} = 1-(1-w)q(\infty). \quad (2{\cdot}7\,\mathrm{A})$$

Let us write

$$\lambda(\phi-c)(\phi-d)$$

for the quadratic which occurs on the right-hand side of the approximating differential equation. We have so far used z and w as the customary complex arguments of generating functions, so that $|z| \leqslant 1$ and $|w| \leqslant 1$, but in fact we can just as well suppose that z and w are

* The letter A attached to an equation number indicates that Assumption (A) was used in the derivation of that equation.

† If a grey cell is still reckoned as part of the grey clone after the cell's 'death', then μ inside the braces in (2·4 A) should be replaced by μz, and corresponding adjustments will be necessary in the following formulae.

real and satisfy $0 \leqslant z < 1$ and $0 \leqslant w < 1$ (and let z and w approach 1 from below when we want to compute moments, etc.). If we do this, then the roots of the characteristic quadratic will be *real* and *distinct* and *positive*. Let us suppose for definiteness that $0 < c < d$; of course c and d will depend on the variable w. The approximating differential equation then has the solution

$$\frac{\phi - c}{\phi - d} = \frac{z - c}{z - d} \exp\left[-(d - c)\lambda t\right],$$

and so when t tends to infinity $\phi(z, w, t)$ (or rather, our approximation) approaches the limit $c(w)$, which is independent of z.

It will be useful to record a few facts about the function $c(w)$. Evidently it is given by

$$c(w) = \frac{1}{2\lambda}\{\lambda + \mu + \nu - \sqrt{[(\lambda + \mu + \nu)^2 - 4\lambda(\mu + \nu\bar{q})]}\}, \tag{2·8 A}$$

where $\bar{q} = 1 - (1 - w)(1 - M/L)$, and it is an elementary matter to expand this in powers of w by the binomial theorem. Notice that (2·8 A) *does* generate a probability distribution, for if $w = 1$ then the square root becomes $(\mu + \nu - \lambda)$ and so $c(1) = 1$; also it is easy to verify that in the power-series expansion all the coefficients are positive. This suggests that our approximate calculation has an exact interpretation, and in fact this is so.

Consider what happens after a long time t has elapsed. The grey clone will have become extinct, and before dying out it will have produced a certain number of 'pioneer' black cells, each of which will have generated a black clone and *some* of which will have generated black clones which themselves escape extinction. These latter black clones are the ones which will be counted by the experimenter (for they will each be very large in size). Accordingly from purely biological arguments we can see that, when $t \to \infty$, the function $\phi(z, w, t)$ will ultimately become independent of t and of z, and will converge to a function of w alone which will generate the distribution of the total number of black clones which ultimately escape extinction. Next we observe that the equation (2·7 A), which we used before as an approximation to (2·5 A), has the following exact interpretation; it is the differential equation which with the boundary condition (2·6 A) determines a generating function for the distribution of the number of grey cells g_t, together with the number of black clones *which were formed during the period* $(0, t)$ and which *ultimately* will escape extinction. It is now clear that, when $t \to \infty$, the solutions to (2·5 A) and (2·7 A) (in each case with (2·6 A) as a side condition) will converge to the same limit. Thus, as an exact result, we have the very useful formula

$$\lim_{t \to \infty} \phi(z, w, t) = c(w) \quad (0 \leqslant z < 1, \quad 0 \leqslant w < 1). \tag{2·9 A}$$

An analytical proof of this would also be possible, starting with the integral equation (2·4 A) and proceeding as in § 4 of Waugh's paper (1955), but the present method is adequate for our purposes and is much more illuminating.

3. The general problem

In the last section we confined attention to a single grey clone; we must now consider the formation of the grey clones and the ensemble of black clones which are derived from them. We start by thinking of the events which follow from the possibility of a grey cell being formed in the time interval $(u, u + du)$, where $0 < u < t$ and $(0, t)$ is the interval of exposure. A grey cell will be formed then with a probability $f(u)\,du$. Suppose that one *is* formed, and

that at the end of the experiment it yields g grey cells and B' *detected* black clones. Suppose that it yields G' *detected* grey clones, where $G' = 0$ or 1. Then

$$E\{z^{G'}w^{B'} \mid g, B'\} = (1-\gamma^g)\,zw^{B'} + \gamma^g w^{B'},$$

and so

$$E\{z^{G'}w^{B'}\} = \phi(\gamma, w, t-u) + z[\phi(1, w, t-u) - \phi(\gamma, w, t-u)], \tag{3·1}$$

and so the contribution to the double generating function for the distribution of the numbers of grey and black clones formed during $(0, t)$ and detected at the epoch t, due to the possibility of a grey cell being formed during $(u, u+du)$, is

$$1 - f(u)\,du + f(u)\,du\,E\{z^{G'}w^{B'}\} + o(du),$$

and to find the generating function for the joint distribution of the total numbers X_t and Y_t of *detected* grey and black clones, we have to form the product of such expressions for each sub-interval $(u, u+du)$ in the interval of observation. Thus, finally, we get

$$\psi(z, w, t) = E\{z^{X_t}w^{Y_t}\} = \exp[-R(w, t) - (1-z)\,S(w, t)], \tag{3·2}$$

where

$$\left.\begin{aligned} R(w, t) &= \int_0^t [1 - \phi(1, w, t-u)]\,f(u)\,du, \\ S(w, t) &= \int_0^t [\phi(1, w, t-u) - \phi(\gamma, w, t-u)]\,f(u)\,du. \end{aligned}\right\} \tag{3·3}$$

The direct use of these formulae is hardly practicable, because we do not have explicit formulae for the function ϕ, but we can obtain useful and interesting results of two quite different types by pursuing the argument a little further.

First, let us calculate the means, variances and covariance of the random variables X_t and Y_t. We find that

$$E\{X_t\} = S(1, t) = \int_0^t [1 - \phi(\gamma, 1, t-u)]\,f(u)\,du; \tag{3·4}$$

$$E\{Y_t\} = -\frac{d}{dw}R(1, t) = \int_0^t \frac{d}{dw}\phi(1, 1, t-u)\,f(u)\,du; \tag{3·5}$$

$$E\{X_t^2\} = [S(1, t)]^2 + S(1, t);$$

$$\operatorname{var}(X_t) = E\{X_t\}; \tag{3·6}$$

$$E\{Y_t^2\} = \left[\frac{d}{dw}R(1, t)\right]^2 - \frac{d^2}{dw^2}R(1, t) - \frac{d}{dw}R(1, t);$$

$$\operatorname{var}(Y_t) = -\left(\frac{d^2}{dw^2} + \frac{d}{dw}\right)R(1, t) = \int_0^t \left[\frac{d^2}{dw^2} + \frac{d}{dw}\right]\phi(1, 1, t-u)\,f(u)\,du; \tag{3·7}$$

$$\operatorname{cov}(X_t, Y_t) = \frac{d}{dw}S(1, t) = \int_0^t \frac{d}{dw}[\phi(1, 1, t-u) - \phi(\gamma, 1, t-u)]\,f(u)\,du. \tag{3·8}$$

In order to use these formulae* we need to know $\phi(\gamma, 1, t)$, the first and second w-derivatives of $\phi(1, w, t)$ at $w = 1$, and the first w-derivative of $\phi(\gamma, w, t)$ at $w = 1$. We can calculate

* In all these formulae the w-differentiations are to be carried out before w is made equal to unity.

these (assuming (A) and the constancy of ν) by going back to equations (2·5–6A). Let us write ϕ', ϕ'' for the first and second w-derivatives of ϕ. Then we shall have

$$D\phi' = 2\lambda\phi\phi' - (\lambda+\mu+\nu)\,\phi' + \nu q(t),$$
$$D\phi'' = 2\lambda\phi\phi'' + 2\lambda\phi'^2 - (\lambda+\mu+\nu)\,\phi'',$$

and

$$\phi(z,1,0) = z, \quad \phi'(z,1,0) = 0, \quad \text{and} \quad \phi''(z,1,0) = 0.$$

Now put $j(t) = \phi(\gamma,1,t)$; $h_1(t) = \phi'(1,1,t)$; $h_2(t) = \phi''(1,1,t)$; and $k(t) = \phi'(\gamma,1,t)$. Then $j(t)$ is the solution to the differential equation

$$Dj = \lambda j^2 - (\lambda+\mu+\nu)\,j + (\mu+\nu), \quad j(0) = \gamma,$$

and so it is given by

$$\frac{j(t)-s}{j(t)-s'} = \frac{\gamma-s}{\gamma-s'}\,e^{-(s'-s)\lambda t},$$

where s and s' are the roots of the quadratic equation

$$\lambda s^2 - (\lambda+\mu+\nu)\,s + (\mu+\nu) = 0;$$

that is, $s = 1$, $s' = (\mu+\nu)/\lambda$, and

$$j(t) = 1 - \left(\frac{\mu+\nu-\lambda}{\lambda}\right)\sum_{m=1}^{\infty}\left(\frac{\lambda(1-\gamma)}{\mu+\nu-\lambda\gamma}\right)^m \exp\left[-m(\mu+\nu-\lambda)\,t\right]. \tag{3·9A}$$

The function $q(t)$ is as before given by (2·2), and $h_1(t)$, $h_2(t)$, and $k(t)$ can now be determined as the solutions to the linear first-order differential equations

$$\left.\begin{aligned} Dh_1 + (\mu+\nu-\lambda)\,h_1 &= \nu q(t), & h_1(0) &= 0;\\ Dh_2 + (\mu+\nu-\lambda)\,h_2 &= 2\lambda h_1^2, & h_2(0) &= 0;\\ Dk + (\mu+\nu+\lambda-2\lambda j)\,k &= \nu q(t), & k(0) &= 0. \end{aligned}\right\} \tag{3·10A}$$

In this way we can determine all the auxiliary functions needed if (3·4)–(3·8) are to be used in the calculation of the first- and second-order moments of (X_t, Y_t).

We now return to the basic formula (3·2) and observe that the marginal distributions of X_t and Y_t are generated by

$$E\{z^{X_t}\} = \exp\left[(z-1)\,S(1,t)\right] \quad E\{w^{Y_t}\} = \exp\left[-R(w,t)\right], \tag{3·11}$$

so that X_t has a Poisson distribution with mean $S(1,t)$. With the usual assumptions, this reduces to

$$E\{X_t\} = \int_0^t [1-j(t-u)]\,f(u)\,du, \tag{3·12A}$$

where $j(t)$ is given by (3·9A).

It is not so easy to draw immediate conclusions from the generating function for the marginal distribution of Y_t, because it depends on a ϕ-function which we have not been able to determine in closed form. However, let us make Assumption (A), let us further assume that the 'feeding' function f is constant throughout the period of exposure $(0,t)$, and let us examine the moments of (X_t, Y_t) when the period of exposure t is large. We then easily obtain the following asymptotic formulae

$$\lim_{t\to\infty} E\{X_t\} = \frac{f}{\lambda}\log\left(\frac{\mu+\nu-\lambda\gamma}{\mu+\nu-\lambda}\right), \tag{3·13A}$$

$$\lim_{t\to\infty} E\{Y_t\}/t = \frac{f\nu(1-M/L)}{\mu+\nu-\lambda}, \tag{3·14A}$$

$$\lim_{t\to\infty} \operatorname{var}(Y_t)/t = \frac{f\nu(1-M/L)}{\mu+\nu-\lambda} + \frac{2\lambda f\nu^2(1-M/L)^2}{(\mu+\nu-\lambda)^3}, \tag{3·15A}$$

$$\lim_{t\to\infty} \operatorname{cov}(X_t, Y_t) = \frac{2\lambda f}{\mu+\nu-\lambda}\int_0^\infty [1-j(\tau)]\,k(\tau)\,d\tau < \infty. \tag{3·16A}$$

Only the last of these results requires any explanation. From (3·8) we know that

$$\operatorname{cov}(X_t, Y_t) = f\int_0^t [h_1(\tau)-k(\tau)]\,d\tau,$$

and from (3·10A) we know that $r(t) = h_1(t)-k(t)$ satisfies the differential equation

$$Dr+(\mu+\nu-\lambda)r = 2\lambda(1-j)\,k, \quad r(0) = 0.$$

Now $j(\infty) = 1$, $k(\infty) = c'(1) = h_1(\infty)$, and so* $r(\infty) = 0$. Thus, integrating the last differential equation, we see that

$$\int_0^\infty r(\tau)\,d\tau = \frac{2\lambda}{\mu+\nu-\lambda}\int_0^\infty [1-j(\tau)]\,k(\tau)\,d\tau,$$

the integral being convergent. It will be noticed that the covariance is positive.

In conclusion it should be noted that all our formulae will require adjustment if grey or black cells are already present in the system at the epoch $t = 0$. It will by now be clear how such adjustments can be made.

4. THE SO-CALLED 'THRESHOLD' EFFECT

Neyman (1958) and Neyman & Scott (1959) are at variance with some earlier writers on the 'two-hit' theory of carcinogenesis, in that they do not agree with the existence of what has (misleadingly) been called a 'threshold'; to be more specific, Neyman and Scott conclude that the effects of the carcinogen are in direct proportion to its intensity. It therefore seems appropriate to offer a comment on this apparent disagreement. The relevant formula is (3·14A), which gives the expected rate of incidence of malignant tumours for continued (long) exposure to a carcinogenic agent maintained at a constant intensity f. (Formula (3·14A) relates in the first instance to the number of *detected* tumours, but it also applies to the *total* number B_t of tumours, as may be seen from the fact that it does not involve the detection-parameter β.) We must now distinguish two possibilities:

(i) The second-order mutations occur at a rate independent of the carcinogen, so that ν is an absolute constant. This is the assumption made by Neyman & Scott, and it makes the incidence of tumours strictly proportional to the intensity f, as they reported in their papers.

(ii) The second-order mutations occur at a rate proportional to the intensity of the carcinogenic agent, so that $\nu = \nu_0 f$, where ν_0 is an absolute constant. Then

$$\lim_{t\to\infty} E\{B_t\}/t = \frac{f^2}{\nu_0 f+(\mu-\lambda)}\nu_0(1-M/L), \tag{4·1}$$

and so now the incidence of tumours is for low intensities (i.e. $f \ll (\mu-\lambda)/\nu_0$) proportional to the intensity *squared*, and it becomes proportional to the first power of the intensity only when the latter is large.

Thus the second assumption implies a difference in the character of the response at low and high intensities, and so it could be said that in this case there is a sort of threshold. But the use of this term is unfortunate because the transition is not at all abrupt.

* It is easily verified that $c'(1)$ is finite.

References

Arley, N. & Iverson, S. (1952). On the mechanism of experimental carcinogenesis. *Acta Path. et Microbiol. Scand.* **31**, 164–71.

Armitage, P. & Doll, R. (1957). A two stage theory of carcinogenesis in relation to the age distribution of human cancer. *Brit. J. Cancer*, **11**, 161–9.

Kendall, D. G. (1952). Les processus stochastiques de croissance en biologie. *Ann. Inst. Poincaré*, **13**, 43–108.

Neyman, J. (1958). A stochastic model of carcinogenesis. (Approximate text of seminar talks delivered at the National Institutes of Health, November 12 and November 26, 1958.)

Neyman, J. & Scott, E. L. (1959). On certain stochastic models of population dynamics. (To be published in *Science*.)

Waugh, W. A. O'N. (1955). An age-dependent birth-and-death process. *Biometrika*, **42**, 291–306.

Biometrika (1960), **47**, 1 *and* 2, *p.* 23
Printed in Great Britain

The quantal response analysis of a series of biological assays on the same subjects

BY J. R. ASHFORD,* C. S. SMITH AND SUSANNAH BROWN
National Coal Board, Pneumoconiosis Field Research

SUMMARY

This paper deals with some problems associated with the analysis of the results of a series of separate biological assays on the same group of subjects, the reaction of the individual organism to the applied stimulus being assumed to be non-reversible. A brief description is given of the methods applied in the analysis of the results of single assays both for single poisons and for mixtures of poisons, with particular reference to the effect of errors of measurement of the observed quantal response. These methods are extended to cover a series of two or more assays, first, under conditions in which the response is measured without error, and secondly, when the response measurement is subject to error. The application of the proposed procedures is illustrated by an example from the field of research in industrial medicine.

1. INTRODUCTION

As a general rule the subjects selected for testing in a biological assay have not previously been exposed to the material under investigation, nor to any other material which affects the same physiological system. When the reaction of the individual organism to the applied stimulus is non-reversible there are, however, some circumstances in which it is convenient to make use of the same subjects for repeated experimentation. Situations of this kind are encountered in studies of human populations, for example, when periodic medical examinations are carried out on groups of men exposed to some health hazard which is a function of time.

This paper is concerned with methods of analysing the results of a series of two or more separate assays on the same group of subjects. In the applications considered each assay involves the same poison or, alternatively, several poisons are applied on each assay, provided they all affect the same physiological system (i.e. under conditions of 'similar joint action'). The reaction of the subject to the applied stimulus is assumed to be non-reversible and consideration is given to the situation in which the observed quantal response is subject to errors of measurement. The effect of such errors is particularly important when dealing with more than one assay, since errors can lead to a response being observed on an earlier assay but not on a later one.

There are numerous references in the literature to methods for analysing the results of biological assays in which the reaction of the individual organism (measured without error) depends on the time elapsing after the poison was administered, as well as on the dose of the poison (Finney, 1952). White & Graca (1958) have described an application of this kind in which each organism was examined at a number of pre-specified times after the poison was applied, and show how the relationship between the probability that the quantal response

* Now with U.K. Atomic Energy Authority, Winfrith, Dorchester, Dorset.

will be observed and the time at which the examination was carried out may be derived. The situation considered in this paper may be regarded as a generalization of that described by White & Graca, in the sense that the effect of time between successive examinations following the initial administration of the poison is equivalent to a special case of the application of a number of doses of another poison in a series of assays on the same subjects. White & Graca do not consider the complications caused by errors in the measurement of the response.

2. THE ANALYSIS OF SINGLE ASSAYS

The statistical methods for analysing quantal response data arising from assays on subjects with no previous exposure to the material under investigation are well established (Finney, 1952). The approach normally adopted is based on the concept of the 'tolerance' of the individual organism, which is defined as the dose which would be just sufficient to produce the quantal response and may be expected to show some variation from subject to subject. It is usual to assume that there exists a transformation, say $f(x)$, of the applied dosage x such that the distribution of tolerances in the population under consideration, denoted $\phi\{f(x)\}$, is symmetrical, with zero mean and unit variance and covering the whole range of values of $f(x)$ between $-\infty$ and $+\infty$. Under these circumstances the probability of response at dose x may be written

$$P = \int_{-\infty}^{f(x)} \phi(\theta)\, d\theta, \tag{1}$$

where $f(x)$ is termed the 'equivalent deviation'. Various mathematical forms have been proposed to represent the tolerance distribution, of which the most commonly used in practical applications are the standard normal and logistic distributions.

When a single poison is applied it is commonly found that the equivalent deviation may be represented by a linear function of the logarithm of the dose, of the form

$$f(x) = a + b \log x, \tag{2}$$

where a and b characterize the response of the population to the particular poison. If the test subjects are exposed to a mixture of different poisons which affect the same physiological system the concept of the tolerance distribution may still be applied. When a mixture of w different poisons $X_1, X_2, \ldots, X_w$ is applied at dose $(x_1, x_2, \ldots, x_w)$, the equivalent deviation is a function of the applied dose, say $f_w(x_1, x_2, \ldots, x_w)$ (Plackett & Hewlett, 1952). If the various poisons making up the mixture do not interact (i.e. under conditions of 'simple similar action') it may be shown (Ashford, 1958*a*) that the equivalent deviation may be written

$$f_w(x_1, x_2, \ldots, x_w) = \theta_w \log_h \left[\sum_{v=1}^{w} h^{(a_v + b_v \log x_v)/\theta_w} \right], \tag{3}$$

where the parameters a_v, b_v characterize the action of the poison X_v when applied singly. If the poisons interact the equivalent deviation will in general be of a more complicated form than (3), depending on the nature of the interaction.

If the test subjects are assigned at random to k groups and the ith group includes n_i subjects, of which r_i manifest the characteristic response when the group is exposed to a mixture of w poisons $X_1, X_2, \ldots, X_w$ applied at dose $(x_{i1}, x_{i2}, \ldots, x_{iw})$, the logarithm of the likelihood of any given set of observations may be written as

$$L = \text{constant} + \sum_{i=1}^{k} [r_i \log P_i + (n_i - r_i) \log (1 - P_i)], \tag{4}$$

where $P_i = \int_{-\infty}^{f_i} \phi(\theta)\, d\theta$ [from (1)] and f_i is a function of the applied dose, and of a number of unknown parameters $u, v, \ldots$. The maximum likelihood estimates of these parameters may be calculated directly or by the solution of a series of equations of the form,

$$\frac{\partial L}{\partial u} = \sum_{i=1}^{k} \frac{n_i(p_i - P_i)}{P_i(1-P_i)} \left(\frac{\partial P_i}{\partial u}\right) = 0, \tag{5}$$

where $p_i = r_i/n_i$ is the observed proportion of responses. The asymptotic variances and covariances of the maximum likelihood estimates may be evaluated in the usual way and result in

$$[\mathrm{cov}\,(\hat{u}, \hat{v})] = \left[-E\left(\frac{\partial^2 L}{\partial u\, \partial v}\right)\right]^{-1} = \left[\sum_{i=1}^{k} \frac{n_i}{P_i(1-P_i)} \left(\frac{\partial P_i}{\partial u}\right)\left(\frac{\partial P_i}{\partial v}\right)\right]^{-1}. \tag{6}$$

The preceding analysis is based on the assumption that the quantal response measurement is not subject to error. There are, however, some circumstances in which this assumption is not justifiable. For example, when the observed quantal response corresponds to a subdivision of an underlying quantitative response scale and the subject is classified as 'responding' or 'not responding' according to whether or not the underlying reaction is in excess of some specified value, it is commonly found that the quantitative measurement is subject to errors, which in turn give rise to corresponding errors in the measurement of the quantal response. In the case of similar joint action it is reasonable to assume that when a mixture of w different poisons is applied at dose $(x_1, x_2, \ldots, x_w)$ the true quantitative response y of any particular subject may be expressed in the form

$$y = z(x_1, x_2, \ldots, x_w) + \epsilon, \tag{7}$$

where ϵ depends only on the subject and is independent of the applied dose. We shall assume further that ϵ is normally distributed in the population of subjects with zero mean and variance σ^2 and that the observed value y' of the quantitative response to any given dose is normally distributed about the true response y, with bias λ and variance ρ^2, where λ and ρ characterize the error distribution. In these circumstances the observed response values y' are normally distributed, with mean $(z+\lambda)$ and variance $(\sigma^2+\rho^2)$. If the characteristic quantal response is observed when $y' \geqslant Y$,* the probability of this event may be written

$$\begin{aligned} P(y' \geqslant Y) &= \int_{Y}^{\infty} \frac{1}{\sqrt{(2\pi(\sigma^2+\rho^2))}} \exp\left\{-\frac{1}{2}\frac{(t-z-\lambda)^2}{\sigma^2+\rho^2}\right\} dt \\ &= \int_{-\infty}^{\left\{\frac{z+\lambda-Y}{\sqrt{(\sigma^2+\rho^2)}}\right\}} \frac{1}{\sqrt{(2\pi)}} \exp\left\{-\tfrac{1}{2}\theta^2\right\} d\theta \\ &= \Phi\left(\frac{z+\lambda-Y}{\sqrt{(\sigma^2+\rho^2)}}\right). \end{aligned} \tag{8}$$

Comparison with equation (1) shows that the equivalent deviation corresponding to the response being observed may be written

$$f_w = \frac{z+\lambda-Y}{\sqrt{(\sigma^2+\rho^2)}}. \tag{9}$$

In the case of a single poison it may be shown (Ashford, 1959*a*) that when the 'true' quantal dosage response relationship is linear the corresponding 'apparent' dosage response relation-

* This approach may be extended to cover the case of a semi-quantal response (i.e. $y < Y_1$, $Y_1 \leqslant y < Y_2$, $y \geqslant Y_2$, etc.). (See Ashford, 1959*a*.)

ship is also linear, although the slope of the regression line is reduced. Equation (9) shows how this conclusion may be extended to mixtures of poisons. Thus, the relationship between the probability of response *being observed* (as opposed to definitely taking place) may be described by an expression of the form (1) and the same methods of analysis may be applied whether or not the response measurement is subject to error. A similar remark applies even when there is no underlying quantitative response, provided the errors are reasonably small and expression (1) holds good to a sufficient degree of approximation.

3. Repeated assays when the response measurement is made without error

It will be assumed that a succession of w separate assays of poisons $X_1, X_2, \ldots, X_w$ [some or all of which may be the same] is carried out on a batch of subjects, the poison applied on the jth assay being denoted X_j. Each poison affects the same physiological system and leads to the same characteristic quantal response, which is non-reversible and is not subject to errors of observation. The test subjects are assigned at random to k groups and the ith group is exposed to a dose x_{ij} of the poison X_j on the jth assay. The total number of subjects in the ith group is denoted n_i and the number who have manifested the characteristic response by the end of the jth assay (i.e. as a result of the 1st, 2nd, ..., and jth assays) is denoted r_{ij} $(1 \leqslant j \leqslant w)$. The number not manifesting the characteristic response after the last wth assay is denoted $s_{i(w+1)}$. As the response measurement is not subject to error it may be asserted that $r_{i(j+1)} \geqslant r_{ij}$ for $j = 1, 2, \ldots, (w-1)$.

On the jth assay the probability of response will depend on the doses of the various poisons applied on the jth and previous assays and, for a subject belonging to the ith group, is given from (1) by the expression

$$P_{ij} = \int_{-\infty}^{f_{ij}} \phi(\theta)\, d\theta, \tag{10}$$

where f_{ij} is a function of the dose $(x_{i1}, x_{i2}, \ldots, x_{ij})$ and of a number of unknown parameters. The probability of response as a result of the jth assay (i.e. no response after the $(j-1)$th assay but response after the jth assay) may be written

$$Q_{ij} = P_{ij} - P_{i(j-1)}, \tag{11}$$

$j = 1, 2, \ldots, w$, and the probability of no response after the last assay may be written

$$Q_{i(w+1)} = 1 - P_{iw}. \tag{12}$$

The number of subjects responding as a result of the jth assay may be written $s_{ij} = [r_{ij} - r_{i(j-1)}]$ and the probability of s_{ij} responses in the ith group as a result of the jth assay $[j = 1, 2, \ldots, w]$ is given by the multinomial distribution, by the expression

$$P(s_{ij}) = \left[n_i! \Big/ \prod_{j=1}^{(w+1)} s_{ij}! \right] \left[\prod_{j=1}^{(w+1)} Q_{ij}^{s_{ij}} \right], \tag{13}$$

as there are $(w+1)$ mutually exclusive possibilities [any subject may respond as a result of the 1st, 2nd, ..., wth assays or may not respond even after the wth assay]. Thus, the logarithm of the likelihood of any particular set of observations may be written

$$L = \text{constant} + \sum_{i=1}^{k} \sum_{j=1}^{(w+1)} s_{ij} \log Q_{ij}. \tag{14}$$

If $u, v \ldots$ denote parameters contained in Q_{ij} the corresponding maximum likelihood estimates are given by the solution of a series of equations of the form

$$\frac{\partial L}{\partial u} = \sum_{i=1}^{k} \sum_{j=1}^{(w+1)} \frac{s_{ij}}{Q_{ij}} \left(\frac{\partial Q_{ij}}{\partial u} \right) = 0. \tag{15}$$

The asymptotic variances and covariances of these estimates are given by the expression

$$[\operatorname{cov}(\hat{u}, \hat{v})] = \left[-E\left(\frac{\partial^2 L}{\partial u \, \partial v} \right) \right]^{-1} = \left[\sum_{i=1}^{k} n_i \sum_{j=1}^{(w+1)} \frac{1}{Q_{ij}} \left(\frac{\partial Q_{ij}}{\partial u} \right) \left(\frac{\partial Q_{ij}}{\partial v} \right) \right]^{-1}. \tag{16}$$

4. Repeated assays when the response measurement is subject to error

If the response measurement is subject to error it is possible that a subject may be classified as showing evidence of response if in fact the response has not taken place. Similarly, the response may not be observed even although it has actually occurred. When a series of assays is considered the fact that a subject may appear to have responded after one particular assay does not necessarily imply that the subject will appear to respond on subsequent assays. The results of each assay must therefore be considered separately and the methods described in the previous section are not applicable. Further treatment of the problem must depend on the assumed form of the error distribution and consideration will be given to the situation described in § 2, which is probably the most important occurring in practical applications. The general methods of approach may, however, be extended to cover other assumptions.

If two successive assays of mixtures of w different poisons (some of which may be at zero dose) are carried out on the same group of subjects, suppose that the resulting true quantitative responses are denoted y_1 and y_2 and the corresponding observed quantitative responses are denoted y_1' and y_2'. On the first assay the doses are $(x_1^{(1)}, x_2^{(1)}, \ldots, x_w^{(1)})$ and that after the second assay the total stimulus applied (on both assays) comprises doses $(x_1^{(2)}, x_2^{(2)}, \ldots, x_w^{(2)})$. Following the notation in § 2, let z_1 and z_2 denote the corresponding values of z on the first and second assays, and let the means of the corresponding error distributions be denoted λ_1, λ_2 and the variances ρ_1^2, ρ_2^2.

As a result of the two assays each subject may be assigned to one of four classes, according to whether or not the quantal response was observed on each assay.

Class 1: $y_1' < Y, y_2' < Y$. No response on either assay.

Class 2: $y_1' < Y, y_2' \geqslant Y$. No response on first assay, but response on second.

Class 3: $y_1' \geqslant Y, y_2' < Y$. Response on first assay, no response on second.

Class 4: $y_1' \geqslant Y, y_2' \geqslant Y$. Response on both assays.

Let $T_{i1}, T_{i2}, T_{i3}, T_{i4}$ denote the probabilities that a man belonging to the ith group is assigned to one of these four classes. If $p(y_1', y_2')$ is the joint probability distribution of y_1' and y_2' for given doses, we have

$$T_{i1} = \int_{-\infty}^{Y} \int_{-\infty}^{Y} p \, dy_1' \, dy_2', \tag{17}$$

$$T_{i2} = \int_{Y}^{\infty} \int_{-\infty}^{Y} p \, dy_1' \, dy_2', \tag{18}$$

$$T_{i3} = \int_{-\infty}^{Y} \int_{Y}^{\infty} p \, dy_1' \, dy_2', \tag{19}$$

$$T_{i4} = \int_{Y}^{\infty} \int_{Y}^{\infty} p \, dy_1' \, dy_2'. \tag{20}$$

Now, from (7) the observed quantitative responses may be expressed in the form

$$y_1' = z_1 + \epsilon + \xi_1 \tag{21}$$

and

$$y_2' = z_2 + \epsilon + \xi_2, \tag{22}$$

where $z_2 \geqslant z_1$, $\epsilon \sim N(0, \sigma^2)$ and $\xi_i \sim N(\lambda_i, \rho_i^2)$, $i = 1, 2$.

Since ϵ, ξ_1, and ξ_2 are independent the joint distribution of (y_1', y_2') is bivariate normal and we have

$$\text{cov}\,(y_1', y_2') = \sigma^2, \quad \text{var}\, y_1' = \sigma^2 + \rho_1^2, \quad \text{var}\, y_2' = \sigma^2 + \rho_2^2.$$

For convenience of presentation we shall assume that the same error distributions apply on both assays, so that

$$\lambda_1 = \lambda_2 = 0 \quad \text{and} \quad \rho_1 = \rho_2 = \rho.$$

The correlation coefficient is thus $t = \sigma^2/(\sigma^2 + \rho^2)$, and the means of y_1, y_2 are z_1 and z_2.

If t_{i1}, t_{i2}, t_{i3} and t_{i4} are respectively the numbers of men in the ith group falling into each of the four classes defined above the likelihood function will take the form

$$L = \sum_{i=1}^{k} (t_{i1} \log T_{i1} + t_{i2} \log T_{i2} + t_{i3} \log T_{i3} + t_{i4} \log T_{i4}). \tag{23}$$

In general, the expression (23) is a very complicated function of σ, ρ and of the parameters contained in z_1 and z_2, but when there are no errors of measurement (i.e. $\rho = 0$) it may readily be shown that it reduces to the form (14).

In practice ρ^2 may be large compared with σ^2. Using a result proved in Kendall (1943, p. 355) the integrals (17)–(20) may be expressed as power series of $t = \sigma^2/(\sigma^2 + \rho^2)$ as follows

$$T_{i4} = \sum_{\nu=0}^{\infty} \Phi^{(\nu)}(U_1)\, \Phi^{(\nu)}(U_2) \frac{t^\nu}{\nu!}, \tag{24}$$

$$T_{i2} = \Phi(U_2) - T_{i4}, \tag{25}$$

$$T_{i3} = \Phi(U_1) - T_{i4}, \tag{26}$$

$$T_{i1} = 1 - \Phi(U_1) - \Phi(U_2) + T_{i4}, \tag{27}$$

where

$$U_j = \frac{z_j - Y}{\sqrt{(\sigma^2 + \rho^2)}} \quad (j = 1, 2).$$

Note that (9) shows that U_j are the values of f_w, the equivalent deviation.

If σ^2/ρ^2 is so small that t may be neglected, we have

$$\begin{aligned} L &= \sum_i [t_{i1} \log (\{1 - \Phi(U_1)\}\{1 - \Phi(U_2)\}) + t_{i2} \log (\Phi(U_2)\{1 - \Phi(U_1)\}) \\ &\qquad + t_{i3} \log (\Phi(U_1)\{1 - \Phi(U_2)\}) + t_{i4} \log \{\Phi(U_1)\, \Phi(U_2)\}] \\ &= \sum_i [(t_{i1} + t_{i2}) \log \{1 - \Phi(U_1)\} + (t_{i1} + t_{i3}) \log \{1 - \Phi(U_2)\} \\ &\qquad + (t_{i2} + t_{i4}) \log \Phi(U_2) + (t_{i3} + t_{i4}) \log \Phi(U_1)], \end{aligned} \tag{28}$$

where $U_j = (z_j - Y)/\rho$.

If σ^2/ρ^2 is not negligible, further terms of the power series (24) must be used, and L must then be regarded as a function of t and of the parameters contained in f_w. The maximum likelihood estimates of these parameters may be obtained in the usual way and the asymptotic variances and covariances of these estimates take the form

$$[\text{cov}\,(\hat{u}, \hat{v})] = \left[-E\left\{\frac{\partial^2 L}{\partial u\, \partial v}\right\}\right]^{-1} = \left[\sum_i n_i \sum_{s=1}^{4} \left\{\frac{1}{T_{is}} \left(\frac{\partial T_{is}}{\partial u}\right)\left(\frac{\partial T_{is}}{\partial v}\right)\right\}\right]^{-1}. \tag{29}$$

This method of approach may be extended to cover the analysis of the results of three or more assays on the same subjects, although the computations involved become progressively more complex.

5. A PRACTICAL APPLICATION

The practical application of the procedures described above may be illustrated by an example from the field of research into the causes of pneumoconiosis amongst coal miners. As part of the National Coal Board's Pneumoconiosis Field Research (Fay, 1957) a series of radiological surveys is being carried out at a number of collieries at intervals of about 4 years. At each of the selected collieries a continuous programme of environmental measurement (Ashford, 1958*b*) provides an estimate of known accuracy of the exposure to airborne dust between successive radiological surveys of each man belonging to the colliery population.

The classification of the X-ray films obtained on the radiological surveys may be regarded for the purposes of this analysis as taking the form of a quantal response—on any particular survey the man may or may not appear to have pneumoconiosis. The film-reading process is subject to errors, which may conveniently be regarded as homoscedastic and normally distributed on an underlying subjective abnormality scale (Ashford, 1959*b*). The development of the condition (as defined in these terms) is believed to be a non-reversible process.

At the time of the first radiological surveys the majority of the men examined have a record of environmental exposure extending back to the time when they joined the mining industry. In general there is no specific information about environmental conditions prior to the first radiological surveys, but it has been found that an indirect measure of previous exposure may be obtained in terms of the period spent in one or more types of general environment (e.g. coal face, coal-getting shift). The individual exposures since the time of the first radiological surveys may be obtained directly, by reference to the results of the systematic programmes of environmental measurement. The man's 'dosage' at the time of the first radiological survey is therefore given by the period spent previously in one or more classes of general environment and his 'dosage' at the time of the second and subsequent surveys consists of two distinct components, corresponding to the 'measured' exposure since the first survey and the previous 'unmeasured' exposure. On general grounds it may be assumed that there can be no interaction between the 'measured' and 'unmeasured' components of exposure and that the effect is the same as that of a mixture of two poisons under conditions of simple similar action, in the conventional terms of biological assay.

After the second medical survey at a colliery the situation is therefore equivalent to two successive assays on the same group of subjects.There are two possible ways in which the quantal response may be measured. The first of these is to classify simultaneously the pair of X-ray films taken of each individual on the two surveys. Because of differences between surveys, due to various causes (such as technical quality), the film reader will be aware of which film was taken on the first survey and which on the second. He will therefore read only 'no response' on both surveys, 'no response' on the first survey and 'response' on the second, or 'response' on both surveys, and will exclude the possibility of 'response' on the first survey but 'no response' on the second. The situation is therefore similar to that considered in § 3, the effect of errors of reading being merely to reduce the slope of the quantal response line. Alternatively, the films obtained on successive surveys may be read independently. In these circumstances the reading of 'response' on the first survey and 'no

response' on the second is not excluded and the situation is similar to that described in § 4.

For the purposes of illustration only the first (side-by-side) method of classification will be considered further. The data obtained at one particular colliery at which this method was applied are given at Table 1.

Table 1. *Prevalence of pneumoconiosis on two successive surveys and corresponding environmental exposure*

			s_{i1}			s_{i2}			s_{i3}		
x_{i1}	x_{i2}	n_i	Obs.		Exp.	Obs.		Exp.	Obs.		Exp.
0·34	500	8	0		0·0	0		0·1	8		7·9
7·00	527	3	0		0·6	0		0·2	3		2·2
32·51	778	6	5		4·5	0		0·2	1		1·3
26·48	1104	8	6	12	5·5 14·7	1	9	0·4 5·7	1	69	2·1 69·6
8·75	1333	4	1		1·1	1		0·6	2		2·3
11·86	1345	7	0		2·7	5		0·9	2		3·4
0·58	1355	54	0		0·3	2		3·4	52		50·3
16·29	1367	7	3		3·5	0		0·8	4		2·7
3·58	1432	12	3		1·1	0		1·7	9		9·3
21·98	1452	5	3	10	3·1 8·6	1	4	0·4 4·5	1	19	1·5 19·9
0·40	2182	5	0		0·0	1		0·6	4		4·3
7·50	2248	4	1		0·9	2		1·0	1		2·1
9·40	3047	5	3		1·5	0		1·5	2		2·0
2·31	3082	8	0		0·4	1		2·5	7		5·1
20·51	3250	2	0	6	1·2 5·8	0	4	0·4 7·3	2	15	0·5 11·8
26·98	3372	4	3		2·8	0		0·5	1		0·7
0·67	4363	6	0		0·0	3		2·4	3		3·5
31·99	4414	2	2		1·5	0		0·3	0		0·3
8·67	4527	3	1		0·8	2		1·2	0		0·9
19·50	4665	2	1		1·1	1		0·5	0		0·4
26·00	4751	1	1	11	0·7 10·0	0	11	0·2 10·2	0	5	0·1 6·8
1·88	5115	6	0		0·2	4		3·1	2		2·7
13·84	5940	13	6		5·7	4		4·9	3		2·4
1·88	6065	45	1		1·5	30		27·0	14		16·5
33·96	6079	4	3		3·0	1		0·6	0		0·4
7·28	6178	11	5		2·5	2		5·8	4		2·7
26·18	6178	5	2	14	3·4 14·2	1	35	1·0 36·0	2	22	0·6 20·9
20·42	6249	5	3		2·9	1		1·4	1		0·7
41·02	6249	1	0		0·8	0		0·1	1		0·1
Total		246	53		53·3	63		63·6	130		129·0

x_{i1} = period (years) on the coal face, coal-getting shift, prior to first survey, weighted means.
x_{i2} = dust exposure between first and second surveys, weighted means.
n_i = total number of men in group.
s_{i1} = number of men with pneumoconiosis of first survey.
s_{i2} = number of men developing pneumoconiosis between first and second surveys.
s_{i3} = number of men without pneumoconiosis on both surveys.

At this colliery the 'unmeasured' environmental exposure prior to the first survey is expressed in terms of the period spent on the coal face, coal-getting shift (denoted x_1) and the 'measured' exposure between the first and second surveys is expressed in terms of arbitrary units of dust exposure (denoted x_2).

Estimates of the five parameters a_1, b_1, a_2, b_2 and θ associated with the response of the population to the 'unmeasured' and 'measured' components of exposure were obtained by the solution of equations (15), by the Newton–Raphson procedure. The tolerance distribution was assumed to be logistic in form and the values of P_{ij} and its partial derivatives were found using the expression,

$$\frac{\partial P_{ij}}{\partial u} = \left(\frac{\partial P_{ij}}{\partial f_{ij}}\right)\left(\frac{\partial f_{ij}}{\partial u}\right) = P_{ij}(1-P_{ij})\left(\frac{\partial f_{ij}}{\partial u}\right), \tag{30}$$

where the u's are the parameters and the f's are given by equations (2) and (3).

The initial approximations for a_1, b_1, a_2, and b_2 were obtained by reference to the marginal distributions of x_1 and x_2. The final values of these estimates were as follows

$$\hat{a}_1 = -4{\cdot}32, \quad \hat{b}_1 = 3{\cdot}57, \quad \hat{a}_2 = -10{\cdot}20, \quad \hat{b}_2 = 5{\cdot}53 \quad \text{and} \quad \hat{\theta} = 5{\cdot}22.$$

The variances of these estimates, obtained by the application of equation (16) are as follows

$$\operatorname{var}\hat{a}_1 = 0{\cdot}361, \quad \operatorname{var}\hat{b}_1 = 0{\cdot}276, \quad \operatorname{var}\hat{a}_2 = 5{\cdot}975,$$
$$\operatorname{var}\hat{b}_2 = 1{\cdot}498, \quad \operatorname{var}\hat{\theta} = 1{\cdot}766.$$

The covariances of corresponding values of $\hat{a}_r$ and $\hat{b}_r$ are

$$\operatorname{cov}(\hat{a}_1, \hat{b}_1) = -0{\cdot}298 \quad \text{and} \quad \operatorname{cov}(\hat{a}_2, \hat{b}_2) = -2{\cdot}928.$$

The hazard relating to either of the two components of exposure may conveniently be expressed in terms of the ED 50—the dose at which 50 % of the population may be expected to manifest the response if the men concerned were exposed only to the given component. If this quantity is denoted d and $D = \log d$, we have, from (2)

$$\hat{D} = \log \hat{d} = -\hat{a}/\hat{b}.$$

From Fieller's theorem (Finney, 1952, p. 29) the 95 % fiducial limits for the ED 50's are found to be

$$\hat{D}_1 = 1{\cdot}210 \pm 0{\cdot}1175, \quad \text{i.e. } \hat{d}_1 = 16 \quad (21 \text{ years}, 12 \text{ years});$$

and

$$\hat{D}_2 = 1{\cdot}844 \pm 0{\cdot}1880, \quad \text{i.e. } \hat{d}_2 = 70 \quad (108 \text{ years}, 45 \text{ years}).$$

The fit of the observations to the hypothesis on which the analysis is based may be examined in Table 1, which gives details of both the 'observed' and 'expected' numbers of men showing evidence of pneumoconiosis. The agreement between the two sets of figures is very good and there is no evidence of any systematic difference between them. The value of χ^2, after grouping certain of the entries for which the expected numbers are small, is 5·777, which for 7 D.F. is not significant ($p \doteqdot 58\,\%$).

Similar data are being obtained at other collieries, both by means of independent and side-by-side readings. The calculations involved are extremely complex and suitable electronic computer programmes are in course of preparation.

References

Ashford, J. R. (1958*a*). Quantal responses to mixtures of poisons under conditions of simple similar action. The analysis of uncontrolled data. *Biometrika*, **45**, 74–88.

Ashford, J. R. (1958*b*). The design of a long term sampling programme to measure the hazard associated with an industrial environment. *J. R. Statist. Soc.* A, **121**, 333–47.

Ashford, J. R. (1959*a*). An approach to the analysis of data for semi-quantal responses in biological assay. *Biometrics*, **15**, 573–81.

ASHFORD, J. R. (1959*b*). A problem of subjective classification in industrial medicine. *Applied Statistics*, **8**, 168–85.

FAY, J. W. J. (1957). The National Coal Board's Pneumoconiosis Field Research. *Nature, Lond.*, **180**, 309–10.

FINNEY, D. J. (1952). *Statistical Method in Biological Assay*. London: Charles Griffin and Co.

KENDALL, M. G. (1943). *The Advanced Theory of Statistics*, **1**. London: Charles Griffin and Co.

PLACKETT, R. L. & HEWLETT, P. S. (1952). Quantal responses to mixtures of poisons. *J. R. Statist. Soc.* B, **14**, 141–63.

WHITE, R. F. & GRACA, J. G. (1958). Multinomially grouped response times for the quantal response bio-assay. *Biometrics*, **14**, 462–88.

Biometrika (1960), **47**, 1 *and* 2, *p.* 33
Printed in Great Britain

Some consequences of superimposed error in time series analysis

BY A. M. WALKER
Statistical Laboratory, University of Cambridge

1. INTRODUCTION

It is well known that time series analysis usually becomes much more complicated when it is necessary to take account of superimposed observational error. However, although it is now many years since attention was first drawn to the problem of superimposed error (this has, for example, been discussed by Kendall (1944) and Quenouille (1947) with particular reference to series generated by a second-order autoregressive process), a detailed study of the complications which can occur does not yet seem to have been made. In this paper some of the problems that might be considered in such a study are formulated, and the difficulties that arise in attempting to solve these are discussed. Unfortunately it would appear that little progress is possible without very laborious algebraic manipulations—this, together with the fact that the mathematics involved is rather uninteresting, may account for the neglect of the topic by writers on the theory of time series analysis. Nevertheless, a few specific results are obtained, and it is hoped that these may stimulate further investigations which should at any rate be of practical importance even if unrewarding in their theoretical aspects.

It will be assumed, as is customary, that the errors are additive and form an independent stationary series which is independent of the original time series. Thus if x_t denotes the observation at time t

$$x_t = u_t + \eta_t, \tag{1}$$

where (η_t) is a completely random series, i.e. a set of independent random variables with a common distribution (whose mean may be taken as zero without loss of generality) and is independent of the stationary series (u_t). It will also be assumed that the values of t are equally spaced, so that they may be taken to be integers without loss of generality, and that (u_t) is generated by a linear autoregressive process, i.e. is the stationary solution of an equation of the form

$$u_t + a_1 u_{t-1} + \ldots + a_p u_{t-p} = v_t, \tag{2}$$

where (v_t) is an independent stationary series, and the constants a_i are such that the moduli of the roots of the equation $z^p + a_1 z^{p-1} + \ldots + a_p = 0$ are less than unity. The condition on (u_t) is less restrictive than might appear at first sight, since by taking p to be sufficiently large one can obtain an adequate approximation to most linear processes, which are such that $u_t = \sum_{r=0}^{\infty} g_r v_{t-r}$, where $\sum_{r=0}^{\infty} g_r^2$ is finite (this is certainly true if $G(z) = \sum_{r=0}^{\infty} g_r z^r$ has radius of convergence greater than unity and has no zeros inside or on the unit circle, so that $[G(z)]^{-1}$ is analytic with a power series expansion for $|z| \leqslant 1$).

Finally, the data will be assumed to consist of values of x for n consecutive values of t, which we denote by $x_1, x_2, \ldots, x_n$.

2. ESTIMATION OF PARAMETERS

From (1) and (2)

$$x_t + a_1 x_{t-1} + \ldots + a_p x_{t-p} = \eta_t + a_1 \eta_{t-1} + \ldots + a_p \eta_{t-p} + v_t$$
$$= y_t, \quad \text{say}, \tag{3}$$

so that (x_t) is generated by a linear autoregressive process with residuals y_t which form a p-dependent stationary process, y_t being independent of y_{t-r} only for values of $r > p$. The correlation between y_t and y_{t-r} $(r \leqslant p)$ is a function of all the parameters $a_1, a_2, \ldots, a_p$, as well as of the ratio $\lambda = \sigma'^2/\sigma^2$, where $\sigma'^2 = \operatorname{var} \eta_t$, $\sigma^2 = \operatorname{var} v_t$; it is therefore not surprising that the problem of estimating $a_1, a_2, \ldots, a_p$ should be much less straightforward than when there is no superimposed error.

Suppose now that the processes (u_t), (η_t) and therefore (x_t) are normal. The behaviour of (x_t) is then completely determined by its autocovariance function $\gamma_x(s) = E(x_t x_{t+s})$ [we may take $E(x_t)$ to be zero without loss of generality], or its spectral density

$$f_x(\omega) = \frac{1}{2\pi} \sum_{s=-\infty}^{\infty} e^{-i\omega s} \gamma_x(s).$$

It is easily seen that

$$2\pi f_x(\omega) = \sigma'^2 + \sigma^2 \{(1 + a_1 e^{i\omega} + \ldots + a_p e^{ip\omega})(1 + a_1 e^{-i\omega} + \ldots + a_p e^{-ip\omega})\}^{-1} \tag{4}$$

and that $$\gamma_x(s) = \gamma_u(s) \quad (s > 0), \qquad \gamma_x(0) = \gamma_u(0) + \sigma'^2,$$

so that the effect of the superimposed error is to reduce all the autocorrelations of the series in the same ratio, multiplying them by the factor $\{1 + (\sigma'^2/\sigma_u^2)\}^{-1}$, where $\sigma_u^2 = \operatorname{var} u_t$.

The most obvious way in which to obtain estimates of the complete set of parameters occurring in the autocovariance function, i.e. $a_1, a_2, \ldots, a_p$, σ^2 (or σ_u^2) and σ'^2, would be to equate the first $p+2$ sample serial covariances $C_s = \sum_{t=1}^{n-s} x_t x_{t+s}/(n-s)$ $(s = 0, 1, \ldots, p+1)$ to their expectations $\gamma_x(s)$. [If $E(x_t)$ is not known *a priori* one may modify these in the usual way, replacing x_t by $x_t - \bar{x}$, where $\bar{x}$ is the sample mean.] This would give estimation equations

$$\left.\begin{aligned} &\frac{C_i}{C_1} = \frac{\hat{\rho}_u(i)}{\hat{\rho}_u(1)} \quad (i = 2, 3, \ldots, p+1), \\ &C_1 = \hat{\sigma}_u^2 \hat{\rho}_u(1), \quad C_0 = \hat{\sigma}_u^2 + \hat{\sigma}'^2, \end{aligned}\right\} \tag{5}$$

where, if $\rho_u(i) = f_i(a_1, a_2, \ldots, a_p)$ is the ith autocorrelation of the series (u_t) expressed as a function of $a_1, a_2, \ldots, a_p$ (which will be the ratio of two polynomials), $\hat{\rho}_u(i) = f_i(\hat{a}_1, \hat{a}_2, \ldots, \hat{a}_p)$.

The estimates obtained from (5) will certainly be consistent and have asymptotic standard errors which are $O(n^{-\frac{1}{2}})$ $(n \to \infty)$. In fact the limiting joint distribution of

$$n^{\frac{1}{2}}(\hat{a}_1 - a_1), \quad n^{\frac{1}{2}}(\hat{a}_2 - a_2), \ \ldots, \ n^{\frac{1}{2}}(\hat{a}_p - a_p), \quad n^{\frac{1}{2}}(\hat{\sigma}^2 - \sigma^2), \quad n^{\frac{1}{2}}(\hat{\sigma}'^2 - \sigma'^2)$$

is multivariate normal with zero means and finite covariance matrix; this follows from the asymptotic normality of $(C_0, C_1, \ldots, C_{p+1})$ which is easily established (see, for example, Walker, 1954).

There is, however, the disadvantage that the method involves the solution of an algebraic equation which appears to be in general very complicated. Thus $\hat{a}_1, \hat{a}_2, \ldots, \hat{a}_p$ are obtained from the first p equations of (5) and since

$$\rho_u(i) + a_1 \rho_u(i-1) + \ldots + a_p \rho_u(i-p) = 0 \quad (i = 2, 3, \ldots, p+1),$$

these are equivalent to the set

$$\left.\begin{aligned} C_{p+1}+\hat{a}_1 C_p+\ldots+\hat{a}_p C_1 &= 0, \\ C_i+\hat{a}_1 C_{i-1}+\ldots+(\hat{a}_i C_1/\hat{\rho}_u(1))+\ldots+\hat{a}_p C_{i-p} &= 0 \quad (i=2,3,\ldots,p) \end{aligned}\right\} \tag{6}$$

(where for $s<0$, C_s is defined as C_{-s}), which determine $\hat{a}_1, \hat{a}_2, \ldots, \hat{a}_p$ as ratios of two polynomials in $[\hat{\rho}_u(1)]^{-1}$. Substitution of these polynomials in $f_1(\hat{a}_1, \hat{a}_2, \ldots, \hat{a}_p)$ will then give an equation for $\hat{\rho}_u(1)$. Of course (6) suggests a simple way of proceeding by successive approximation, $\hat{a}_1, \hat{a}_2, \ldots, \hat{a}_p$ being calculated using the value of $\hat{\rho}_u(1)$ determined by the previous approximation, but the complicated form of the equation makes it very difficult, for example, to evaluate the asymptotic variances of the estimates.

Other simpler methods which make use of some serial covariances for lags greater than $p+1$ have therefore been suggested. For example, in the case of a second-order autoregressive process, $\hat{a}_1$ and $\hat{a}_2$ could be obtained from

$$C_3+\hat{a}_1 C_2+\hat{a}_2 C_1 = 0, \quad C_4+\hat{a}_1 C_3+\hat{a}_2 C_2 = 0,$$

as pointed out by Quenouille (1947, p. 127).

Again for general p, since

$$\sum_{i=0}^{p} a_i \rho_x(r-i) = 0 \quad (r>p)$$

(defining a_0 as unity), we can use any p equations of the form

$$\sum_{i=0}^{p} \hat{a}_i C_{r-i} = 0 \tag{7}$$

(see, for example, Bartlett, 1955, p. 265). As $E(C_s) \to 0$ when $s \to \infty$ one would in general expect to obtain more accurate estimates from $r = p+1, p+2, \ldots, 2p$ than from other sets of p values of r.

Having thus determined $\hat{a}_1, \hat{a}_2, \ldots, \hat{a}_p$ we obtain $\hat{\sigma}_u^2$ and $\hat{\sigma}'^2$ as before from the last two equations of (5).

This brings us to the first main problem, the relative and absolute asymptotic efficiencies of the various methods of estimation. Let the first method, using serial covariances $C_0, C_1, \ldots, C_{p+1}$, be referred to as Method A, and the second method, using equations of the form (7), as Method B. Then the following questions may be asked:

(i) What is the best choice of values of r in Method B, judging this by the asymptotic efficiency of the estimates?

(ii) How does the efficiency† of Method B using these values (or any other convenient set such as $r = p+m, p+m+1, \ldots, 2p+m-1$ $(m>0)$) compare with that of Method A?

(iii) What can be said about the absolute efficiencies of Methods A and B?

To these we may add:

(iv) How far are the answers to (i) and (ii) affected by non-normality of (u_t) or (η_t)?

(for (iii) normality of (u_t) and (η_t) is clearly vital).

The importance of the normality assumption, which of course in practice will often not be realistic, may be regarded as the second main problem. One could of course consider other methods of estimation, such as those using all the information contained in the serial covariances $C_0, C_1, \ldots, C_k$, where k is an arbitrary integer exceeding $p+1$, but in view of the difficulties that arise in dealing adequately with Methods A and B this would hardly seem to be worth while.

† Throughout this paper the term 'efficiency' will always mean 'asymptotic efficiency'.

3. EFFICIENCY OF METHODS OF ESTIMATION

All these sets of estimates are easily seen to have similar asymptotic distributional properties—thus $n^{\frac{1}{2}}(\hat{a}_1 - a_1)$, $n^{\frac{1}{2}}(\hat{a}_2 - a_2)$, etc., have in every case a limiting multinormal distribution with zero means. Hence to investigate their efficiencies it is merely necessary to calculate their asymptotic variances (and possibly covariances if some overall measure of efficiency for the whole set is required). In principle this is quite straightforward since one has merely to express the estimates asymptotically as linear functions of the C_i or the sample serial correlations $r_i = C_i/C_0$ and then use the formulae for asymptotic variances and covariances of the C_i or r_i (see, for example, Bartlett, 1955, pp. 255–6). However, the calculations seem to be very tedious.

A fairly detailed examination of the case $p = 1$ was made. Here the situation is particularly simple, since Method A then becomes identical to Method B with $r = 1, 2$, and (6) reduces to the single equation

$$C_2 + C_1\hat{a} = 0 \tag{8}$$

(for convenience we replace a_1 by a in this case). The asymptotic variance of $\hat{a}$ is probably most easily obtained by using the general result

$$\operatorname{cov}(H_+^t C_l, H_+^u C_u) \sim \frac{1}{n}\left\{\text{coefficient of } z^{t-u} \text{ in } G_x^2(z)\left(\sum_{i=0}^{p} a_i z^i\right)\left(\sum_{i=0}^{p} a_i z^{-i}\right)\right\} \quad (t \geqslant u > p), \tag{9}$$

where $H_+ = \sum_{i=0}^{p} a_i E_t^{-i}$, E_t being the forward shift operator that changes t into $t+1$, and $G_x(z)$ the covariance-generating function $\sum_{s=-\infty}^{\infty} \gamma_x(s) z^s$ of (x_t) (see, for example, Bartlett & Diananda, 1950, p. 112). With $p = 1$, $t = u = 2$ we get

$$n \operatorname{var}(C_2 + aC_1) \sim \text{constant term in } G_x^2(z)(1+az)(1+az^{-1}), \tag{10}$$

and from (8) $$\delta(C_2 + C_1 a) \sim -\gamma_x(1)\,\delta\hat{a}$$

(using the standard notation for statistical differentials), so that

$$\operatorname{var}\hat{a} \sim \operatorname{var}(C_2 + C_1 a)/\gamma_x^2(1).$$

Now $\gamma_x(1) = -a$ and

$$G_x(z) = \sigma^2\left\{\lambda + \frac{1}{(1+az)(1+az^{-1})}\right\}$$

$$= \sigma^2\left\{\lambda + \frac{1}{1-a^2}\left(\frac{1}{1+az} - \frac{az^{-1}}{1+az^{-1}}\right)\right\}.$$

Hence, after a little straightforward algebra we obtain

$$n \operatorname{var}\hat{a} \sim \frac{1-a^2}{a^2}\{1 + 2\lambda(1-a^2) + \lambda^2(1-a^4)\}. \tag{11}$$

In general, Method B will give $C_r + \hat{a}C_{r-1} = 0$ (where $r \geqslant 2$). It can be shown similarly that then

$$n \operatorname{var}\hat{a} \sim \frac{1-a^2}{a^{2(r-1)}}\{1 + 2\lambda(1-a^2) + \lambda^2(1-a^4)\}.$$

Hence we may conclude that it is best to take $r = 2$.

To obtain the absolute asymptotic efficiency of $\hat{a}$ we use a result due to Whittle (see, for example, Appendix 2 to *A Study in the Analysis of Stationary Time Series*, by H. Wold,

1954, p. 214), which may be conveniently stated for general p. Asymptotically efficient estimates of $a_1, a_2, \ldots, a_p$ and λ are given by the minimization of $S = n \sum_{s=-\phi(n)}^{\phi(n)} \alpha_s C_s$, where α_s is the coefficient of $e^{i\omega s}$ in the Fourier expansion of $v\{2\pi f_x(\omega)\}^{-1} = \{g(\omega)\}^{-1}$ say. v, the 'residual prediction variance', is given by the constant term in the Laurent expansion of $\log 2\pi f_x(\omega)$ in positive and negative powers of $e^{i\omega}$, i.e.

$$\exp \frac{1}{2\pi} \int_{-\pi}^{\pi} \log 2\pi f_x(\omega)\, d\omega, \quad \text{and} \quad \phi(n) = o(n) \quad (n \to \infty).$$

Thus if

$$1 + \lambda(1 + a_1 e^{i\omega} + \ldots + a_p e^{ip\omega})(1 + a_1 e^{-i\omega} + \ldots + a_p e^{-ip\omega}) \equiv (b_0 + b_1 e^{i\omega} + \ldots + b_p e^{ip\omega})(b_0 + b_1 e^{-i\omega} + \ldots + b_p e^{-ip\omega}),$$

where $b_0, b_1, \ldots, b_p$ are such that the roots of the equation $b_0 + b_1 z + \ldots + b_p z^p = 0$ have moduli greater than unity, then $v = b_0^2 \sigma^2$ and

$$\{g(\omega)\}^{-1} = \frac{b_0^2 (1 + a_1 e^{i\omega} + \ldots + a_p e^{ip\omega})(1 + a_1 e^{-i\omega} + \ldots + a_p e^{-ip\omega})}{(b_0 + b_1 e^{i\omega} + \ldots + b_p e^{ip\omega})(b_0 + b_1 e^{-i\omega} + \ldots + b_p e^{-ip\omega})},$$

the α_s being determined by putting $\{g(\omega)\}^{-1}$ into partial fractions and expanding these. The asymptotic covariance matrix of the estimates $(\hat{a}_1, \hat{a}_2, \ldots, \hat{a}_p, \hat{\lambda}) \equiv (\hat{\theta}_1, \hat{\theta}_2, \ldots, \hat{\theta}_{p+1})$ say, is then V^{-1}, where $V = (v_{ij})$ with

$$v_{ij} = \tfrac{1}{2} n \left\{ \text{constant term in expansion of } \frac{\partial \log g}{\partial \theta_i} \frac{\partial \log g}{\partial \theta_j} \right\}.$$

Taking $p = 1$ we find that the asymptotic variance of the efficient estimate a^* is given by

$$n \operatorname{var} a^* \sim \frac{(1 - a^2)(1 + a\mu_1)^2}{(a + \mu_1)^2}, \tag{12}$$

where μ_1 is the root of the equation

$$\lambda a(z^2 + 1) + z\{1 + \lambda(1 + a^2)\} = 0$$

whose modulus is less than unity. The algebra involved in deriving (12) is somewhat heavy, owing to the fact that b_0 cannot be expressed simply in terms of a and λ. Details of this are given in the Appendix.

The asymptotic efficiency of $\hat{a} = -C_2/C_1$ is therefore

$$E_{\hat{a}} = \frac{a^2(1 + a\mu_1)^2}{(a + \mu_1)^2 [1 + 2\lambda(1 - a^2) + \lambda^2(1 - a^4)]}. \tag{13}$$

Table 1 gives values of $E_{\hat{a}}$ (which is an even function of a) for various values of $|a|$ and λ. From this we see that the efficiency is near unity only when λ or a is small. It is in fact easy to show that $E_{\hat{a}} \to 1$ as $\lambda \to 0$ for fixed $a \neq 0$ or as $a \to 0$ for fixed $\lambda \neq 0$. For large λ, $E_{\hat{a}}$ is approximately $(1 - a^2)^3/(1 + a^2)$.

Table 1. *Values of* $E_{\hat{a}}(a, \lambda)$

	$\lvert a \rvert$				
λ	0·1	0·3	0·5	0·7	0·9
0·5	1·00	0·96	0·91	0·87	0·90
1·0	0·99	0·92	0·80	0·71	0·74
2·0	0·98	0·86	0·66	0·49	0·47
10·0	0·97	0·74	0·43	0·18	0·07

There remains the question of the efficiency of the estimation of the other parameters by Method A. In dealing with this it is convenient to consider λ and $\sigma_x^2 = \sigma_u^2 + \sigma'^2$ rather than σ_u^2 and σ'^2. Here the calculations are even heavier, but it was thought worth while to obtain an explicit formula for the asymptotic efficiency of $\hat{\lambda}$ (it seems likely that an interval estimate would be required for λ more often than for σ_x^2, for example, because it is desired to test the hypothesis $\lambda = 0$, i.e. that no superimposed error is present).

The last two equations of (5) give

$$r_1 = -\hat{a}/\{1+\hat{\lambda}(1-\hat{a}^2)\}.$$

Hence
$$\hat{\lambda} = (r_2 - r_1^2)/(r_1^2 - r_2^2)$$

and
$$\delta\hat{\lambda} \sim \frac{1+\lambda(1-a^2)}{a^2(1-a^2)}[(\delta r_2 + a\delta r_1)\{1+\lambda(1+a^2)\} + a\delta r_1\{1+\lambda(1-a^2)\}]. \tag{14}$$

Now it is easily shown that

$$n\sigma_x^4 \operatorname{var} r_1 \sim (1+2\rho_1^2)\nu_0 - 4\rho_1\nu_1 + \nu_2,$$

$$n\sigma_x^4 \operatorname{cov}(r_1, r_2+ar_1) \sim a\nu_0 + (1-2a\rho_1-a^2)\nu_1 - (a+2\rho_1)\nu_2$$

and
$$n\sigma_x^4 \operatorname{var}(r_2+ar_1) \sim (1+a^2)\nu_0 + 2a\nu_1,$$

(compare equation (10)), where $\rho_1 = \rho_x(1) = -a/\{1+\lambda(1-a^2)\}$ and ν_k is the coefficient of z^k in $G_x^2(z)$. Thus

$$\frac{a^4(1-a^2)^2\sigma_x^4}{\{1+\lambda(1-a^2)\}^2}(n \operatorname{var}\hat{\lambda}) \sim A_0\nu_0 + A_1\nu_1 + A_2\nu_2, \tag{15}$$

where

$$A_0 = (1+a^2)\{1+\lambda(1+a^2)\}^2 + 2a^2\{1+\lambda(1+a^2)\}\{1+\lambda(1-a^2)\} + (1+2\rho_1^2)a^2\{1+\lambda(1-a^2)\}^2,$$
$$A_1 = 2a\{1+\lambda(1+a^2)\}^2 + 2a(1-2a\rho_1-a^2)\{1+\lambda(1+a^2)\}\{1+\lambda(1-a^2)\} - 4\rho_1 a^2\{1+\lambda(1-a^2)\}^2,$$
$$A_2 = -2a(a+2\rho_1)\{1+\lambda(1+a^2)\}\{1+\lambda(1-a^2)\} + a^2\{1+\lambda(1-a^2)\}^2.$$

Now
$$\sigma_x^2 = \sigma^2\{1+\lambda(1-a^2)\}/(1-a^2)$$

and
$$\nu_0 = \frac{\sigma^4}{(1-a^2)^3}\{\lambda^2(1-a^2)^3 + 2\lambda(1-a^2)^2 + 1 + a^2\},$$

$$\nu_s = \frac{\sigma^4(-a)^s}{(1-a^2)^3}\{2\lambda(1-a^2)^2 + s + 1 - (s-1)a^2\} \quad (s > 0).$$

Hence, on expressing the right-hand side of (15) as a quartic in λ, we find

$$a^4(n \operatorname{var}\hat{\lambda}) \sim 1 + 4\lambda(1+a^2) + \lambda^2(6+14a^2+a^4) + 2\lambda^3(2+8a^2+a^4-a^6) + \lambda^4(1+6a^2+a^4). \tag{16}$$

The asymptotic variance of the efficient estimate λ^* obtained by minimizing S is given by

$$\mu_1^4(n \operatorname{var}\lambda^*) \sim \lambda^4(1+a\mu_1)\{(3a-4a^3)\mu_1^5 + \mu_1^4 + 2a\mu_1^3 + (2-4a^2)\mu_1^2 - a\mu_1 + 1\} \tag{17}$$

(this formula is derived in the Appendix).

The asymptotic efficiency of $\hat{\lambda}$ is therefore

$$E_{\hat{\lambda}} = \frac{a^4\lambda^4\{(3a-4a^3)\mu_1^5 + \mu_1^4 + 2a\mu_1^3 + (2-4a^2)\mu_1^2 - a\mu_1 + 1\}(1+a\mu_1)}{\mu_1^4\{1+4\lambda(1+a^2) + \lambda^2(6+14a^2+a^4) + 2\lambda^3(2+8a^2+a^4-a^6) + \lambda^4(1+6a^2+a^4)\}}. \tag{18}$$

No way of simplifying this complicated expression could be discovered.

Values of $E_{\hat{\lambda}}$ are given in Table 2. The efficiency of $\hat{\lambda}$ behaves in the same way as that of $\hat{a}$, being near unity when λ or a is small, and having unity as a limit when $\lambda \to 0$ with a fixed or when $a \to 0$ with λ fixed. For large λ, $E_{\hat{\lambda}}$ is approximately

$$\frac{(1-a^2)^3(1+5a^2+4a^4)}{1+6a^2+a^4}.$$

From Tables 1 and 2 we see that for a wide range of values of a and λ, $\hat{a}$ and $\hat{\lambda}$ are of reasonably high efficiency. The ease with which they may be computed is a considerable practical advantage and if the efficiencies corresponding to $a = \hat{a}, \lambda = \hat{\lambda}$ are both not much less than unity, it would usually not be worth while obtaining estimates by some more complicated method which makes use of the information contained in the sample serial correlations for lags greater than 2. Computation of the efficient estimates λ^* and a^* seems to be more laborious than one would expect. The formulae for partial derivatives of α_s with respect to a and λ are very complicated and the obvious procedure of obtaining successive approximations by Newton's method, taking for example $(\hat{a}, \hat{\lambda})$ as a first approximation, is therefore somewhat cumbersome—in practice it would probably be a good deal quicker to proceed by computing S for a number of values (a, λ) and determining its minimum by some method which does not require any knowledge of its functional form.

Table 2. *Values of* $E_{\hat{\lambda}}(a, \lambda)$

	$\lvert a\rvert$				
λ	0·1	0·3	0·5	0·7	0·9
0·5	1·00	0·96	0·91	0·84	0·77
1·0	0·99	0·92	0·81	0·69	0·56
2·0	0·98	0·86	0·69	0·51	0·35
10·0	0·97	0·76	0·49	0·24	0·08

It may be noted that the conclusions regarding the estimation of a are applicable also to certain series, not subject to superimposed error, which are generated by a first-order linear autoregressive process with residuals forming a 1-dependent stationary process and therefore representable by a two-term moving average of a series of mutually uncorrelated random variables with a common variance. For then

$$x_t + a x_{t-1} = v_t + b v_{t-1}$$

with $$\operatorname{var} v_t = \sigma^2, \quad \operatorname{cov}(v_t, v_{t+s}) = 0 \quad (s > 0),$$

giving $$2\pi f_x(\omega) = \frac{\sigma^2(1+b\,e^{i\omega})(1+b\,e^{-i\omega})}{(1+a\,e^{i\omega})(1+a\,e^{-i\omega})}. \tag{19}$$

Thus the spectral density (which determines the correlation structure) is of the same general form as (4) [§ 2] with $p = 1$. However, taking $|b| \leqslant 1$ (which gives the 'regular' moving average representation), we see that for (19) to reduce to (4), $-b = \mu_1$, the root of $\lambda a(z^2+1) + z\{1+\lambda(1+a^2)\} = 0$ with modulus less than unity. Now μ_1 clearly lies between 0 and $-a$ (the value corresponding to $\lambda = \infty$), so that for the reduction to be possible, $0 \leqslant b \leqslant a$ $(a > 0)$ or $0 \geqslant b \geqslant a$ $(a < 0)$. The efficiency of estimation of both a and b for *all* normal series with spectral densities of the form (19) ($|b| \leqslant 1$, $|a| < 1$), using the above method could of course be investigated in exactly the same way; it seems likely that the calculations would be equally tedious.

The general estimation problem for processes with spectral density given by the analogue of (19) for $p > 1$, viz.

$$2\pi f_x(\omega) = \frac{\sigma^2(1+b_1e^{i\omega}+\ldots+b_pe^{ip\omega})(1+b_1e^{-i\omega}+\ldots+b_pe^{-ip\omega})}{(1+a_1e^{i\omega}+\ldots+a_pe^{ip\omega})(1+a_1e^{-i\omega}+\ldots+a_pe^{-ip\omega})},$$

is incidentally quite different from that for the process with spectral density given by (4) because of the greater number of parameters involved ($2p+1$ compared with $p+2$).

A similar examination of the case $p = 2$ would at any rate be practicable, although the work involved would be very great and tables of values of efficiencies would be of triple entry. Here, moreover, Method A is fairly straightforward since it gives

$$C_3+\hat{a}_1C_2+\hat{a}_2C_1 = 0,$$
$$C_2+\hat{a}_1C_1+(\hat{a}_2C_1/\hat{\rho}_u(1)) = 0,$$

with $\hat{\rho}_u(1) = -\hat{a}_1/(1+\hat{a}_2)$, so that it requires the solution only of a quadratic equation, and this may be written as an equation for $\hat{a}_1$, namely

$$\hat{a}_1^2(C_1^2-C_2^2)+2\hat{a}_1C_2(C_1-C_3)+C_3(C_1-C_3) = 0,$$

(see Quenouille, 1947, p. 126, equation (7)).

For $p > 2$, however, a general investigation would be a most formidable undertaking, and probably quite impracticable. Nevertheless, one might hope that then some useful information on efficiency would be obtained by considering particular estimates for various values of $a_1, a_2, \ldots, \lambda$.

4. Effect of non-normality

If superimposed error were not present, the effect of non-normality would not be at all serious. For (x_t) would be a linear process and so the asymptotic distributional properties of the sample serial correlations r_i would not be affected, the joint distribution of $n^{\frac{1}{2}}(r_i-\rho_i)$ $(i = 1, 2, 3, \ldots)$ remaining asymptotically normal with zero means and limiting covariance matrix with elements given by Bartlett's formulae [see, for example, Bartlett (1955), p. 256]. Thus the asymptotic behaviour of estimates $\hat{a}_1, \hat{a}_2, \ldots, \hat{a}_p$ obtained from equations of the form (7) with $r = 1, 2, \ldots, p$ (which are asymptotically efficient under the normality assumption) would remain the same [as can in fact be shown directly, see Bartlett (1955), pp. 258–9]. The asymptotic variance of $C_0 = \hat{\sigma}_x^2$ would of course contain an additional term (a multiple of the fourth cumulant of x) but very often we would not want to estimate this. But with superimposed error (x_t) becomes the sum of *two* independent linear processes which in general is certainly not a linear process. Now if (x_t) is of this form, so that we can write

$$x_t = \sum_r (g_r y_{t-r}+h_r z_{t-r}),$$

where (y_t) and (z_t) are independent completely random series, the contribution of the fourth cumulant terms to $\lim_{n\to\infty} n \operatorname{cov}(C_i, C_j)$ is

$$\sum_{v=-\infty}^{\infty} \kappa(x_t, x_{t+v}, x_{t+i}, x_{t+v+i+j}) = \kappa_4(y)\, G_iG_j+\kappa_4(z)\, H_iH_j, \tag{20}$$

where $\kappa_4(y)$, $\kappa_4(z)$ are the fourth cumulants of (y_t), (z_t), $G_i = \sum_r g_r g_{r+i}$ and $H_i = \sum_r h_r h_{r+i}$. From (20), it is easily shown that the contribution to $\lim_{n\to\infty} n \operatorname{cov}(r_i, r_j)$ is

$$\{\kappa_2(y)\, G_0+\kappa_2(z)\, H_0\}^{-4}\{\kappa_4(z)\,\kappa_2^2(y)+\kappa_4(y)\,\kappa_2^2(z)\}\{G_iH_0-H_iG_0\}\{G_jH_0-H_jG_0\} \tag{21}$$
$$(\kappa_2(y) = \operatorname{var} y, \quad \kappa_2(z) = \operatorname{var} z).$$

This will not be zero for all i, j unless $G_i = cH_i\,(i = 0, 1, 2, \ldots)$ (which in general will not hold except when $g_r \propto h_r$ so that the process reduces to a single linear one) or

$$\gamma_2(y) + \gamma_2(z) = 0, \tag{22}$$

where γ_2 denotes the standardized measure of flatness κ_4/κ_2^2. It follows that for the superimposed error model the formulae for asymptotic variances and covariances of estimates of $a_1, a_2, \ldots, a_p, \lambda$ obtained by any of the methods that have been discussed (Method A, Method B, or Whittle's asymptotically efficient method) will remain valid in the non-normal case only when (22) holds (there is no difficulty about their asymptotic normality as the asymptotic normality of the sample serial covariances is easily established by the method used in Walker (1954)). However, it seems unlikely that (22) would be satisfied approximately in any practical situation unless *both* the original series and the superimposed error are nearly normal. Hence if an appreciable degree of non-normality is suspected one would always want to take account of the fourth cumulant terms in the asymptotic variance formulae. In principle there is no difficulty in doing this, but it makes the algebra still more troublesome, and the details have not been worked out even for the case $p = 1$. To obtain consistent estimates of these terms one will of course require consistent estimates of $\kappa_4(u)$ and $\kappa_4(\eta)$, but it is fairly easy to see that these can be found in terms of

$$\sum_{t=1}^{n} x_t^4/n, \quad \sum_{t=1}^{n-s} x_t^2 x_{t+s}^2/(n-s) \quad \text{for some } s > 0,$$

and the serial covariances C_i.

It is perhaps worth mentioning that the question of the effect of non-normality in the presence of superimposed error was suggested to the author by a definite practical problem, the analysis of a series of observations of the variation of latitude, which is described by Walker & Young (1955). Here one had essentially a bivariate autoregressive model with superimposed error, $\{x_t, y_t\}$ being the stationary solution of

$$x_t - \alpha x_{t-1} + \beta y_{t-1} = \eta_{1,t} + v_{1,t} - \alpha v_{1,t-1} + \beta v_{2,t-1},$$
$$y_t - \beta x_{t-1} - \alpha y_{t-1} = \eta_{2,t} + v_{2,t} - \beta v_{1,t-1} - \alpha v_{2,t-1},$$

where $(v_{1,t})$, $(v_{2,t})$ are independent completely random series representing the basic disturbances, and $(\eta_{1,t})$, $(\eta_{2,t})$ independent completely random series representing the superimposed observational error. The main problem was the estimation of parameters κ, γ such that $\alpha = e^{-\kappa}\cos\gamma$, $\beta = e^{-\kappa}\sin\gamma$. Estimates $\hat{\kappa}$ and $\hat{\gamma}$ were obtained using the sample serial autocovariances and cross-covariances for lags 1 and 2 and asymptotic variance formulae derived on the assumption that all the above completely random series were normal. There were reasonable physical grounds for supposing the v series to be approximately normal but there was some doubt about the normality of the η series, although this was not sufficient to make it seem worth while undertaking the laborious task of modifying the variance formulae.

I wish to thank Mr D. A. East for the computation of Tables 1 and 2.

REFERENCES

BARTLETT, M. S. (1955). *An Introduction to Stochastic Processes.* Cambridge University Press.
BARTLETT, M. S. & DIANANDA, P. H. (1950). *J. Roy. Statist. Soc.* B, **12**, 108–15.
KENDALL, M. G. (1944). *Biometrika*, **33**, 105–22.
QUENOUILLE, M. H. (1947). *J. Roy. Statist. Soc.* A, **110**, 123–9.
WALKER, A. M. (1954). *Proc. Camb. Phil. Soc.* **50**, 60–4.
WALKER, A. M. & YOUNG, A. (1955). *Monthly Not. Roy. Astr. Soc.* **115**, 443–59.
WOLD, H. (1954). *A Study in the Analysis of Stationary Time Series* (2nd edition). Uppsala.

APPENDIX

The formulae for the asymptotic variances of the efficient estimates a^* and λ^* may be obtained as follows.

We have
$$\{g(\omega)\}^{-1} = \frac{b_0^2(1+a\,e^{i\omega})(1+a\,e^{-i\omega})}{1+\lambda(1+a\,e^{i\omega})(1+a\,e^{-i\omega})}. \tag{23}$$

Now the denominator of (23) is equal to
$$\lambda a z^{-1}(z-\mu_1)(z-\mu_2), \quad \text{where} \quad z = e^{i\omega},$$
μ_1 and $\mu_2 = \mu_1^{-1}$ being the roots of the equation
$$\lambda a(z^2+1)+z\{1+\lambda(1+a^2)\} = 0.$$
This is also equal to $(b_0+b_1 z)(b_0+b_1 z^{-1})$ by definition, so that $b_0^2 = -\lambda a\mu_2$. Hence putting (23) into partial fractions gives
$$\{g(\omega)\}^{-1} = -a\mu_2\left\{1-\frac{1}{\lambda a(\mu_1-\mu_2)}\left(\frac{\mu_1}{z-\mu_1}-\frac{\mu_2}{z-\mu_2}\right)\right\}$$
$$= -a\mu_2+\frac{\mu_2}{\lambda(\mu_1-\mu_2)}\{\mu_1 z^{-1}(1-\mu_1 z^{-1})^{-1}+(1-\mu_1 z)^{-1}\}.$$

From this, incidentally, the coefficients of the sample serial covariances in S can easily be seen to be
$$\alpha_0 = (1+a^2+2a\mu_1)/(1-\mu_1^2),$$
and
$$\alpha_s = \alpha_{-s} = \mu_1^{s-1}\{a(1+\mu_1^2)+\mu_1(1+a^2)\}/(1-\mu_1^2) \quad (s>0).$$

Now
$$-\frac{\partial \log g}{\partial \lambda} = \frac{\partial \log b_0^2}{\partial \lambda} - \frac{g^{-1}}{b_0^2}$$
$$= \frac{1}{\lambda^2 a(\mu_1-\mu_2)}\{\mu_1 z^{-1}(1-\mu_1 z^{-1})^{-1}+\mu_1 z(1-\mu_1 z)^{-1}\} \tag{24}$$
since the constant term in its expansion in positive and negative powers of z must vanish.

Similarly
$$-\frac{\partial \log g}{\partial a} = \frac{\partial \log b_0^2}{\partial a}+\frac{z}{1+az}+\frac{z^{-1}}{1+az^{-1}}-\frac{\lambda(z^2+2az+1)}{\lambda a(z^2+1)+z\{1+\lambda(1+a^2)\}}$$
$$= \frac{\partial \log b_0^2}{\partial a}+\frac{z}{1+az}+\frac{z^{-1}}{1+az^{-1}}-\frac{1}{a}+\frac{1+\lambda(1-a^2)}{a^2\lambda(\mu_1-\mu_2)}\left(\frac{\mu_1}{z-\mu_1}-\frac{\mu_2}{z-\mu_2}\right)$$
$$= z(1+az)^{-1}+z^{-1}(1+az^{-1})^{-1}+\frac{1+\lambda(1-a^2)}{a^2\lambda(\mu_1-\mu_2)}\{\mu_1 z^{-1}(1-\mu_1 z^{-1})^{-1}+\mu_1 z(1-\mu_1 z)^{-1}\}, \tag{25}$$
since the constant term in its expansion also vanishes.

The constant terms in the expansions of
$$\{\mu_1 z^{-1}(1-\mu_1 z^{-1})^{-1}+\mu_1 z(1-\mu_1 z)^{-1}\}^2, \quad \{z(1+az)^{-1}+z^{-1}(1+az^{-1})^{-1}\}^2$$
and
$$\{\mu_1 z^{-1}(1-\mu_1 z^{-1})^{-1}+\mu_1 z(1-\mu_1 z)^{-1}\}\{z(1+az)^{-1}+z^{-1}(1+az^{-1})^{-1}\}$$
are easily seen to be respectively
$$2\mu_1^2/(1-\mu_1^2), \quad 2/(1-a^2) \quad \text{and} \quad 2\mu_1/(1+a\mu_1).$$

Hence from (24) and (25), if we write

$$2V/n = U = \begin{pmatrix} u_{11} & u_{12} \\ u_{12} & u_{22} \end{pmatrix},$$

where, for example,

$$u_{11} = \text{constant term in the expansion of } \left(\frac{\partial \log g}{\partial \lambda}\right)^2,$$

then

$$u_{11} = \left(\frac{\mu_1^2}{\lambda^4 a^2 (1-\mu_1^2)^2}\right)\left(\frac{2\mu_1^2}{1-\mu_1^2}\right) = 2\mu_1^4/\lambda^4 a^2 (1-\mu_1^2)^3,$$

$$u_{12} = \frac{1}{\lambda^2 a(\mu_1-\mu_2)}\left\{\frac{2\mu_1}{1+a\mu_1} + \frac{1+\lambda(1-a^2)}{a^2\lambda(\mu_1-\mu_2)}\left(\frac{2\mu_1^2}{1-\mu_1^2}\right)\right\}$$

$$= \frac{2\mu_1^2}{\lambda^2 a(\mu_1^2-1)}\left\{\left(\frac{1}{1+a\mu_1}\right) + \frac{\mu_1(1+2a\mu_1+\mu_1^2)}{a(1-\mu_1^2)^2}\right\}$$

on eliminating λ from the second term in the bracket by means of the relation

$$\lambda\{a\mu_1^2 + (1+a^2)\mu_1 + a\} = -\mu_1,$$

and

$$u_{22} = \frac{2}{1-a^2} + \frac{1+\lambda(1-a^2)}{a^2\lambda(\mu_1-\mu_2)}\left(\frac{4\mu_1}{1+a\mu_1}\right) + \left(\frac{1+\lambda(1-a^2)}{a^2\lambda(\mu_1-\mu_2)}\right)^2 \frac{2\mu_1^2}{1-\mu_1^2}$$

$$= \frac{2}{1-a^2} + \frac{2\mu_1(1+2a\mu_1+\mu_1^2)}{a(1-\mu_1^2)}\left\{\left(\frac{2}{1+a\mu_1}\right) + \frac{\mu_1(1+2a\mu_1+\mu_1^2)}{a(1-\mu_1^2)^2}\right\}.$$

Now if

$$P = (1+a\mu_1)^{-1}, \quad Q = \mu_1(1+2a\mu_1+\mu_1^2)/a(1-\mu_1^2)^2,$$

$$\Delta = \det U = 4\left\{\frac{\mu_1^4}{\lambda^4 a^2(\mu_1^2-1)^3(1-a^2)} + \frac{\mu_1^4}{\lambda^4 a^2(1-\mu_1^2)^2}(2PQ+Q^2-(P+Q)^2)\right\}$$

$$= \frac{4\mu_1^4(a+\mu_1)^2}{\lambda^4 a^2(1-a^2)(1+a\mu_1)^2(1-\mu_1^2)^3}.$$

Hence

$$n \operatorname{var} a^* \sim 2u_{11}/\Delta = \frac{(1-a^2)(1+a\mu_1)^2}{(a+\mu_1)^2},$$

and

$$n \operatorname{var} \lambda^* \sim 2u_{22}/\Delta = \frac{\lambda^4 a^2(1-\mu_1^2)^3(1-a^2)(1+a\mu_1)^2}{\mu_1^4(a+\mu_1)^2}\left\{\frac{1}{1-a^2} + (1-\mu_1^2)Q(2P+Q)\right\},$$

which, on simplifying the expression in brackets, reduces to

$$\lambda^4(1+a\mu_1)\{(3a-4a^3)\mu_1^5 + \mu_1^4 + 2a\mu_1^3 + (2-4a^2)\mu_1^2 - a\mu_1 + 1\}/\mu_1^4.$$

Biometrika (1960), **47**, 1 *and* 2, *p.* 45
Printed in Great Britain

Deterministic customer impatience in the queueing system GI/M/1

BY P. D. FINCH
Research Techniques Division, London School of Economics

1. INTRODUCTION

In this paper the following single server queueing system is considered. Customers arrive at the instants $t_1, t_2, \ldots, t_n, \ldots$ such that the $\tau_n = t_n - t_{n-1}$, $n = 1, 2, \ldots$, $t_0 = 0$ are independently and identically distributed non-negative random variables with common distribution function $A(x)$ and finite expectation $\alpha = \int_0^\infty x\,dA(x)$. Customers are served in the order of their arrival and if a customer arrives to find the server idle then he commences service immediately. A customer will wait for service in the queue only for a time not exceeding a fixed time W. If a customer waits as long as W then he departs never to return. It is supposed that the service times of customers who receive service are independently and identically distributed random variables, independent of the input process $\{t_n\}$ and with common distribution function $B(x) = 1 - e^{-\mu x}$ for $x \geqslant 0$. The case when the input process is a Poisson process, that is, $A(x) = 1 - e^{-\lambda x}$, for $x \geqslant 0$, has been studied by Barrer (1957) who obtained the limiting distribution of queue size at an arbitrary instant of time. Let w_n, $n = 1, 2, \ldots$, denote the time the nth customer waits in the queue, if he departs impatiently then $w_n = W$, if he does not then $w_n < W$. Let $u_n = s_n - \tau_{n+1}$ where s_n is independent of τ_{n+1} and has the exponential distribution of service time and τ_n defined above, has the distribution $A(x)$ of inter-arrival time. Let $U(x)$ be the common distribution function of the identically distributed random variables u_n. Then

$$\left.\begin{aligned} U(x) &= \mu e^{-\mu x}\int_0^\infty A(y)\,e^{+\mu y}\,dy \quad \text{if} \quad x > 0 \\ &= \mu e^{-\mu x}\int_{-x}^\infty A(y)\,e^{+\mu y}\,dy \quad \text{if} \quad x < 0. \end{aligned}\right\} \tag{1}$$

By relating w_{n+1} to w_n in a manner analogous to that of Lindley (1952) and taking into account the characteristic property of the exponential distribution we obtain

$$\left.\begin{aligned} w_{n+1} &= 0 \qquad\quad \text{if} \quad w_n + u_n < 0 \\ &= w_n + u_n \quad \text{if} \quad 0 \leqslant w_n + u_n < W \\ &= W \qquad\quad \text{if} \quad W \leqslant w_n + u_n. \end{aligned}\right\} \tag{2}$$

Write $F_n(x) = P(w_n \leqslant x)$ then it follows from (2) that

$$\left.\begin{aligned} F_{n+1}(x) &= 0 \qquad\qquad\qquad\quad \text{if} \quad x < 0 \\ &= \int_{-\infty}^{x} F_n(x-y)\,dU(y) \quad \text{if} \quad 0 \leqslant x < W \\ &= 1 \qquad\qquad\qquad\quad \text{if} \quad W \leqslant x. \end{aligned}\right\} \tag{3}$$

The sequence of random variables $\{w_n\}$ forms a Markov chain with state space the finite interval $(0, W)$. It follows from the general theory of Markov chains that a limiting distribution $F(x) = \lim_{n\to\infty} F_n(x)$ exists and is independent of the initial distribution $F_1(x)$. It follows from (3) by Lebesgue's convergence theorem that $F(x)$ is the solution to the following integral equation

$$\left.\begin{aligned} F(x) &= 0 && \text{if} \quad x < 0 \\ &= \int_{-\infty}^{x} F(x-y)\,dU(y) && \text{if} \quad 0 \leqslant x < W \\ &= 1 && \text{if} \quad W \leqslant x. \end{aligned}\right\} \tag{4}$$

The limiting distribution $F(x)$ is the unique solution to (4) because if there were a second solution $F^*(x)$ we could take it for the initial solution $F_1(x)$. From (3) we would have $F_n(x) = F^*(x)$ for each $n = 1, 2, \ldots$ and hence $F^*(x) = F(x)$.

Introduce random variables $c_n = W - w_n$ and $v_n = -u_n$ then we have

$$\left.\begin{aligned} c_{n+1} &= 0 && \text{if} \quad c_n + v_n \leqslant 0 \\ &= c_n + v_n && \text{if} \quad 0 < c_n + v_n \leqslant W \\ &= W && \text{if} \quad W < c_n + v_n. \end{aligned}\right\} \tag{5}$$

Write $G_n(x) = P(c_n < x) = 1 - F_n(W-x)$. Then $G(x) = \lim_{n\to\infty} G_n(x)$ exists and

$$G(x) = 1 - F(W-x).$$

Further $G(x)$ is the unique solution to the following integral equation

$$\left.\begin{aligned} G(x) &= 0 && \text{if} \quad x \leqslant 0 \\ &= \int_{-\infty}^{x} G(x-y)\,dV(y) && \text{if} \quad 0 < x \leqslant W \\ &= 1 && \text{if} \quad W < x, \end{aligned}\right\} \tag{6}$$

where $V(x)$ is the common distribution function of the identically distributed random variables v_n, and is given by

$$\left.\begin{aligned} V(x) &= \mu e^{\mu x} \int_{x}^{\infty} A(y)\, e^{-\mu y}\,dy && \text{if} \quad x > 0 \\ &= \mu e^{\mu x} \int_{0}^{\infty} A(y)\, e^{-\mu y}\,dy && \text{if} \quad x < 0. \end{aligned}\right\} \tag{7}$$

The main result of this paper is the solution of the integral equation (6) when $V(x)$ is given by (7).

2. Solution of the integral equation

In this section we prove the following theorem.

Theorem 1. If $G(W, x)$ is the solution to the integral equation (6) when $V(x)$ is given by (7) then

$$G(W, x) = [\mu M(x) - M'(x)][\mu M(W)]^{-1} \quad \text{for} \quad 0 \leqslant x < W, \tag{8}$$

where $M(x)$ is a function whose Laplace transform $m(p) = \int_0^\infty e^{-px} M(x)\,dx$ is given by

$$m(p) = \frac{p-\mu+\mu\{a(p)-a(\mu)\}}{(p-\mu)\{\mu a(p)+p-\mu\}} \tag{9}$$

and

$$a(p) = \int_0^\infty e^{-px}\,dA(x).$$

Proof. The solution $G(W,x)$ to the integral equation (6) is a function of two variables W and x. It is reasonable therefore to try to find a solution of the form $G(W,x) = G_1(W)G_2(x)$. If a $G(W,x)$ of this form is substituted into equation (6) and use made of (7) it is not difficult, though rather lengthy, to determine the functions $G_1(W)$ and $G_2(x)$. It will be found that the solution is of the form (8). In order to prove the theorem here we note that since the solution to (6) exists and is unique it is sufficient to substitute expression (8) into the equation and determine the unknown function $M(x)$.

Writing $N(x) = \mu M(x) - M'(x)$ and performing this substitution we obtain

$$\begin{aligned} N(x) - \int_0^x N(x-y)\,dV(y) + \mu a(\mu)\int_0^x N(x-y)\,e^{\mu y}\,dy \\ = \mu a(\mu)\,e^{\mu x}\left\{M(W)\,e^{-\mu W} + \int_0^W N(u)\,e^{-\mu u}\,du\right\} \\ = \mu a(\mu)\,M(0)\,e^{\mu x}. \end{aligned} \tag{10}$$

This equation has been derived under the assumption that $0 \leqslant x < W$, but W is arbitrary and the equation is independent of W; thus equation (10) is true for all non-negative values of x.

Introduce the Laplace transforms

$$n(p) = \int_0^\infty e^{-px} N(x)\,dx, \quad V^+(p) = \int_0^\infty e^{-px}\,dV(x)$$

then from (10) we obtain

$$n(p)\left[1 - V^+(p) + \frac{\mu a(\mu)}{p-\mu}\right] = \frac{\mu a(\mu)}{p-\mu} M(0). \tag{11}$$

From (7) we have $V^+(p) = \mu\{a(\mu)-a(p)\}/p-\mu$, and hence from (11) we obtain

$$n(p) = \mu a(\mu)\,M(0)/\{p-\mu+\mu a(p)\}. \tag{12}$$

Since $N(x) = \mu M(x) - M'(x)$ we have

$$m(p) = \frac{M(0)}{p-\mu} - \frac{n(p)}{p-\mu}. \tag{13}$$

It is clear from (8) that we need only determine $M(x)$ up to a multiplicative constant, thus we can take $M(0) = 1$. Equation (9) then follows from (12) and (13).

Example. Suppose that the input process $\{t_n\}$ is an Erlang E_k process with mean k/λ, that is $dA(x) = \lambda^k x^{k-1} e^{-\lambda x}\,dx/(k-1)!$ for $x \geqslant 0$. Then $a(p) = \lambda/(\lambda+p)^k$ and from (12)

$$n(p) = \frac{\mu a(\mu)\,(\lambda+p)^k}{(\lambda+p)^k\,(p-\mu)+\mu\lambda^k}. \tag{14}$$

Let $\beta_0, \beta_1, \ldots, \beta_k$ ($\beta_0 = 0$) be the $(k+1)$-roots of $(\lambda+p)^k(p-\mu)+\mu\lambda^k = 0$ (it is easy to see that these roots are in general distinct). Expanding (14) in partial fractions we have

$$n(p) = \sum_{r=0}^{k} \frac{K_r}{p-\beta_r}, \tag{15}$$

where $$K_r = \left[\frac{(p-\beta_r)\mu a(\mu)(\lambda+p)^k}{(\lambda+p)^k(p-\mu)+\mu\lambda^k}\right]_{p=\beta_r} = \frac{\lambda+\beta_r+k(\beta_r-\mu)}{\mu a(\mu)(\lambda+\beta_r)}.$$

From (14) and (15) we obtain $$\sum_{r=0}^{k} \frac{K_r}{\mu-\beta_r} = 1. \tag{16}$$

Taking inverse transforms of (15) we have

$$N(x) = K_0 + \sum_{r=1}^{k} K_r e^{\beta_r x}. \tag{17}$$

But $M(x) = e^{\mu x} - e^{\mu x}\int_0^x e^{-\mu y} N(y)\,dy$ and so taking (16) into account we obtain

$$M(x) = \frac{K_0}{\mu} + \sum_{r=1}^{k} \frac{K_r}{\mu-\beta_r} e^{\beta_r x}. \tag{18}$$

Thus from (8) the solution to (6) with $V(x)$ given by (7) is

$$G(W,x) = \left[K_0 + \sum_{r=1}^{k} K_r e^{\beta_r x}\right]\left[K_0 + \sum_{s=1}^{k} \frac{\mu K_s}{\mu-\beta_s} e^{\beta_s W}\right]^{-1} \quad (0 < x < W) \tag{19}$$

and the limiting distribution of waiting time $F(W,x) = 1 - G(W, W-x)$ is given by

$$F(W,x) = \left[\sum_{r=1}^{k} K_r\left\{\frac{\mu}{\mu-\beta} - e^{-\beta_r x}\right\} e^{\beta_r W}\right]\left[K_0 + \sum_{s=1}^{k} \frac{\mu K_s}{\mu-\beta_s} e^{\beta_s W}\right]^{-1}. \tag{20}$$

In the particular case $k = 1$, that is the input process is a Poisson process, we obtain

$$\left.\begin{aligned} F(W,x) &= \left[1 - \frac{\lambda}{\mu} e^{-(\mu-\lambda)x}\right]\left[1 - \frac{\lambda^2}{\mu^2} e^{-(\mu-\lambda)W}\right]^{-1} \quad (\lambda \neq \mu,\ 0 \leqslant x < W) \\ &= \frac{1+\mu x}{2+\mu W} \qquad\qquad (\lambda = \mu,\ 0 < x < W). \end{aligned}\right\} \tag{21}$$

3. Relation to the case $W = \infty$

If the traffic intensity $\rho = (\mu\alpha)^{-1} < 1$ and $W = \infty$, then it follows from a theorem of Lindley (Lindley, 1952) that a limiting distribution of waiting time, $F(x)$, exists and is the unique solution to (4) with $W = \infty$, in this case the solution to (6) with $W = \infty$ does not exist. When $\rho < 1$, $W = \infty$ and $V(x)$ is given by (7), that is service time has an exponential distribution parameter μ, it is known (Kendall, 1953) that this limiting distribution $F(x)$ is of the following form

$$F(x) = 1 - \pi e^{-\mu(1-\pi)x}, \tag{22}$$

where π is the unique solution in $(0, 1)$ of the equation

$$\pi = \int_0^\infty e^{-(1-\pi)\mu x}\,dA(x).$$

We shall show in a heuristic way that when $\rho < 1$, Theorem 1 gives a solution $F(x)$ of the form (22) when $W \to \infty$.

If $F(W,x) = 1-G(W, W-x)$ is the waiting time distribution when $\rho < 1$ then from (8) we have

$$\mu M(W) F(W,x) = \mu M(W) - \mu M(W-x) + M'(W-x) \quad (0 < x < W). \tag{23}$$

In particular, assuming that $M(x)$, $M'(x)$ are continuous and letting $x \to 0+$ in (23) we obtain

$$\mu M(W) F(W,0) = M'(W). \tag{24}$$

But W is arbitrary and hence equation (24) is true for all $W > 0$. Replacing W by $(W-x)$ in (24) and substituting for $M'(W-x)$ in (23) gives

$$F(W,x) = 1 - \frac{M(W-x)}{M(W)}\{1-F(W-x,0)\}. \tag{25}$$

Let us assume that $\lim_{W\to\infty} F(W,x)$ exists and write $\lim_{W\to\infty} F(W,0) = 1-\theta$ where we assume that $0 < \theta < 1$. (This is not the case if $\rho \geqslant 1$.) From (24) we obtain

$$\lim_{W\to\infty} \frac{M'(W)}{M(W)} = \mu(1-\theta). \tag{26}$$

From (25) and the existence of $\lim_{W\to\infty} F(W,x)$ we deduce that

$$\lim_{W\to\infty} \frac{M(W-x)}{M(W)}$$

exists and that

$$\lim_{W\to\infty} F(W,x) = 1-\theta \lim_{W\to\infty} \frac{M(W-x)}{M(W)}. \tag{27}$$

Write $R(x) = \lim_{W\to\infty} M(W-x)/M(W)$ and assume that it is possible to interchange the limiting operation and differentiation with respect to x. We obtain

$$\begin{aligned} R'(x) &= -\lim_{W\to\infty} \frac{M'(W-x)}{M(W)} \\ &= -\left[\lim_{W\to\infty} \frac{M'(W-x)}{M(W-x)}\right]\left[\lim_{W\to\infty} \frac{M(W-x)}{M(W)}\right] \\ &= -\mu(1-\theta) R(x) \end{aligned}$$

using equation (26). But $R(0) = 1$ and hence $R(x) = e^{-\mu(1-\theta)x}$. Substitution into (27) gives

$$\lim_{W\to\infty} F(W,x) = 1-\theta e^{-\mu(1-\theta)x} \tag{28}$$

which is of the form (22). In order to prove the equivalence of (22) and (28) it would be necessary to show that $\theta = \pi$. We shall not do so here.

If $\rho = 1$ then neither (4) nor (6) possesses a solution when $W = \infty$. If $\rho > 1$ it follows from Lindley's theorem that a solution to (6) with $W = \infty$ exists, but that the solution to (4) with $W = \infty$ does not exist. The case $\rho > 1$ is of interest because the solution to (6) with $V(x)$ given by (7) and $W = \infty$ is the waiting time distribution of a queue with Poisson input process, parameter μ and service time distribution $A(x)$.* For such a queue there exists the well-known Pollaczek–Khintchine formula for the Laplace transform of the waiting

* That is of the dual queueing process (cf. Foster, 1959). The relationship of the duality principle to the queueing process of this paper is not investigated here.

time distribution (Lindley, 1952). We shall now show in a heuristic manner how the Pollaczek–Khintchine formula can be deduced from Theorem 1 when $\rho > 1$. Let us assume that $\lim_{W\to\infty} M(W)$ exists when $\rho > 1$. (This is not the case if $\rho < 1$.) Then

$$\lim_{W\to\infty} M(W) = \lim_{p\to 0} pm(p). \tag{29}$$

Equation (29) is valid if and only if $\int_0^t u\,dM(u) = o(t)$ as $t \to \infty$ (Widder, 1946, Chapter V, Theorem 3.6). This condition is implied by the existence of $\lim_{W\to\infty} M(W)$. Using equation (9) we obtain

$$\lim_{W\to\infty} M(W) = \frac{a(\mu)}{1-\sigma},$$

where $\sigma = \alpha\mu = \rho^{-1}$ and hence from (8)

$$G(x) = \lim_{W\to\infty} G(W,x) = \frac{1-\sigma}{\mu a(\mu)}\{\mu M(x) - M'(x)\}. \tag{30}$$

Write $g(p) = \int_0^\infty e^{-px} G(x)\,dx$. Since $\mu M(x) - M'(x) = N(x)$ with Laplace transform $n(p)$ given by (12), we have $g(p) = (1-\sigma)[p-\mu+\mu a(p)]^{-1}$.

Thus

$$\begin{aligned}\int_{0-}^\infty e^{-px}\,dG(x) &= pg(p)\\ &= (1-\sigma)\left[1-\frac{\mu[1-a(p)]}{p}\right]^{-1}\end{aligned} \tag{31}$$

which is the Pollaczek–Khintchine formula.

It would be of interest to relate the solution to (4) and (6) to the corresponding solutions when they exist, with $W = \infty$. If $\rho < 1$ there does not seem to be any simple relationship, but if $\rho > 1$ we have the following theorem.

THEOREM 2. If $\rho = \{\mu\alpha\}^{-1} > 1$, and if $G(x)$, $G(W,x)$ are the solutions to (6) in the case $W = \infty$, $W < \infty$, respectively, when $V(x)$ is given by (7) then

$$G(W,x) = L(W)\,G(x) \quad (0 < x < W), \tag{32}$$

where

$$L(W) = \mu\, e^{-\mu W}\left[\int_W^\infty G(y)\, e^{-\mu y}\,dy\right]^{-1}, \tag{33}$$

and $G(x)$ has the Laplace–Stieltjes transform (31).

Proof. Substituting (32) into (6) we obtain

$$L(W)\,G(x) = V(x-W) + L(W)\int_{x-W}^x G(x-y)\,dV(y) \quad (0 < x < W).$$

But

$$G(x) = \int_{-\infty}^{x-W} G(x-y)\,dV(y) + \int_{x-W}^x G(x-y)\,dV(y) \quad (x > 0).$$

Hence

$$V(x-W) = L(W)\int_{-\infty}^{x-W} G(x-y)\,dV(y) \quad (0 < x < W).$$

Substituting from (7) we obtain (33).

4. The limiting distribution of queue size

In this section we consider the limiting distribution of queue size at the instants a service commences. It is not difficult to write down formulae for the distribution of queue size at other instants of time, for example, at an arrival or a departure (impatiently or otherwise) but these formulae are very complicated and do not seem to lead to simple expressions even in the case of a Poisson input process.

Let P_n be the limiting probability that n customers are left in the queue at the instant a service commences. P_n is the conditional probability of n arrivals in the waiting time of the customer starting service given that this waiting time does not exceed W. Thus

$$P_0 = F_C(W, 0+) + \int_{0+}^{W} \{1 - A(y)\}\, dF_C(W, y),$$

$$P_n = \int_{0+}^{W} \{A_n(y) - A_{n+1}(y)\}\, dF_C(W, y),$$

where
$$F_C(W, x) = \frac{F(W, x)}{F(W, W-0)} \quad (0 < x < W)$$

and $F(W, x)$ is the solution to (4) and $A_n(x)$ is the n-fold iterated convolution of $A(x)$ with itself.

In the case the input process is Poisson with parameter λ we obtain the following simple expressions

$$P_0 = (1-\rho)(1+\rho\Gamma_0)\{1-\rho e^{-(\mu-\lambda)W}\}^{-1},$$

$$P_n = (1-\rho)\rho^{n+1}\Gamma_n\{1-\rho e^{-(\mu-\lambda)W}\}^{-1} \quad (n \geqslant 1),$$

where $\rho = \lambda/\mu$ and

$$\Gamma_n = 1 - e^{-\mu W}\left[1 + \frac{(\mu W)}{1!} + \ldots + \frac{(\mu W)^n}{n!}\right].$$

5. The many server queue

In this section we discuss in a heuristic way how the results of the previous sections are modified if there are N independent and identical serves ($N > 1$).

Let m_n denote the random variable which specifies the number of customers in the system on the nth arrival. If $m_n < N$ then the nth arrival does not wait and if $m_n < N-1$ then neither the nth nor the $(n+1)$th waits for service. If we suppose that $m_n \geqslant N-1$ then it can be shown that equation (2) remains valid if the random variable s_n has an exponential distribution with parameter μN instead of μ. The meaning of w_n, w_{n+1} in this equation is modified, however, for they are now waiting times conditional upon $m_n \geqslant N-1$. To emphasize this we rewrite the equation in the following way

$$\left.\begin{aligned} (w_{n+1} \mid m_n \geqslant N-1) &= 0 && \text{if} \quad (w_n \mid m_n \geqslant N-1) + u_n < 0 \\ &= (w_n \mid m_n \geqslant N-1) + u_n && \text{if} \quad 0 < (w_n \mid m_n \geqslant N-1) + u_n < W \\ &= W && \text{if} \quad W < (w_n \mid m_n \geqslant N-1) + u_n. \end{aligned}\right\} \quad (34)$$

In the ordinary notation of conditional probabilities we have

$$P(w_{n+1} \leqslant x \mid m_n \geqslant N-1) = P(w_n + u_n \leqslant x \mid m_n \geqslant N-1) \quad (0 < x < W).$$

Let us assume that the limiting distributions for w_n and m_n exist, we prove the following theorem.

THEOREM 3. The limiting distribution $\lim_{n\to\infty} P(w_n \leqslant x \mid m_n \geqslant N-1)$ is the solution to (4) with $U(x)$ given by (1) and μ replaced by $N\mu$.

In order to prove the theorem we require the following lemma.

LEMMA 1.

$$\lim_{n\to\infty} P(w_{n+1} \leqslant x \mid m_{n+1} \geqslant N-1) = \lim_{n\to\infty} P(w_{n+1} \leqslant x \mid m_n \geqslant N-1).$$

We prove the lemma by noting that for $x > 0$

$$P(w_{n+1} \leqslant x) = P(m_{n+1} < N-1) + P(m_{n+1} \geqslant N-1)\, P(w_{n+1} \leqslant x \mid m_{n+1} \geqslant N-1)$$

and $$P(w_{n+1} \leqslant x) = P(m_n < N-1) + P(m_n \geqslant N-1)\, P(w_{n+1} \leqslant x \mid m_n \geqslant N-1),$$

since $w_{n+1} = 0$ whenever m_{n+1} or m_n does not exceed $(N-1)$.

Assuming the existence of limiting distributions for w_{n+1} and m_n the lemma follows at once from these two equations.

It follows from the lemma and equations (34) that the limiting distribution

$$\lim_{n\to\infty} P(w_n \leqslant x \mid m_n \geqslant N-1)$$

satisfies the integral equation (4) with $U(x)$ given by (1) and μ replaced by $N\mu$. This proves the theorem. Finally, we remark that if $F_N(W, x)$ denotes this limiting distribution then the limiting conditional distribution for waiting time given that the customer does wait is just $[F_N(W, x) - F_N(W, 0+)]/[1 - F_N(W, 0+)]$ in virtue of the equation

$$P(w_n \leqslant x \mid w_n > 0,\, m_n \geqslant N-1)\, P(w_n > 0 \mid m_n \geqslant N-1) + P(w_n = 0 \mid m_n \geqslant N-1)$$
$$= P(w_n \leqslant x \mid m_n \geqslant N-1).$$

I am grateful to Dr F. G. Foster for reading an earlier version of this paper and suggesting some substantial improvements in presentation.

REFERENCES

BARRER, D. Y. (1957). Queueing with impatient customers and ordered service. *Opns. Res.* **5**, 650–6.

FOSTER, F. G. (1959). A unified theory for stock storage and queue control. *Operat. Res. Quart.* **10**, 121–30.

KENDALL, D. G. (1953). Stochastic processes occurring in the theory of queues and their analysis by the method of the imbedded Markov chain. *Ann. Math. Statist.* **24**, 338–54.

LINDLEY, D. V. (1952). The theory of queues with a single server. *Proc. Camb. Phil. Soc.* **48**, 277–89.

WIDDER, D. W. (1946). *The Laplace Transform*. Princeton University Press.

Biometrika (1960), **47**, 1 *and* 2, *p.* 53
Printed in Great Britain

The polykays of the natural numbers

BY D. E. BARTON, F. N. DAVID AND EVELYN FIX*
University College London and University of California, Berkeley

It was pointed out by David (1956) that problems concerning the null hypothesis distribution of rank criteria for the random sequence can be simplified if these problems are translated into those of samples drawn without replacement from a finite population. Thus, suppose a random sequence of N elements, n of which possess a given characteristic, and let the elements be numbered $1, 2, \ldots, N$. Any criterion based on the ranks of the n elements can be treated as a criterion based on n elements randomly drawn without replacement from the finite population of the first N natural numbers. This idea leads to a unification of the study of such criteria, and moments of the sampling distribution of any criterion based on symmetric functions of the n ranks may be written down with comparative ease with the aid of theory which has already been set out in general fashion by Wishart (1952). Further, the expressions given by Wishart can, for the case of the finite population of the first N integers, be considerably shortened by expressing them in terms of the generalized Bernoulli numbers. We give here formulae whereby this can be done.

NOTATION

N. E. Nörlund (1920, section 32) defined the Bernoulli polynomial of νth degree and nth order and showed it was the coefficient of $t^\nu/\nu!$ in the development of

$$\frac{t^n e^{xt}}{(e^t-1)^n} = \sum_{\nu=0}^{\infty} B_\nu^{(n)}(x)\frac{t^\nu}{\nu!}.$$

The Bernoulli polynomials of the first order, i.e. when $n = 1$ are those usually referred to as the Bernoulli polynomials. When $n = 1$ and $x = 0$ the numbers $B_\nu^{(1)}$ are those called the Bernoulli numbers by some authors. Others place these numbers in modulus, while others still omit the zero elements. We shall use Nörlund's notation $B_\nu^{(n)}(x)$, writing $B_\nu^{(n)}(0)$ as $B_\nu^{(n)}$. If $B_\nu^{(n)}$ is multiplied by a function on the right we shall incorporate this in square braces to distinguish it from a polynominal argument.

GENERATING FUNCTIONS FOR THE MOMENTS AND CUMULANTS OF THE FIRST N NATURAL NUMBERS

The probability generating function, assuming all N numbers to have equal probability, is

$$P(t) = \frac{t(1-t^N)}{N(1-t)}.$$

From this we may deduce the central moment generating function, the cumulant generating function, and the factorial moment and factorial cumulant generating functions. These are

$$\mathscr{M}(t) = \frac{\sinh(\frac{1}{2}Nt)}{N\sinh(\frac{1}{2}t)}, \qquad \mathscr{K}(t) = \log \mathscr{M}(t),$$

$$\mathscr{F}(t) = \frac{(1+t)^{N+1}-(1+t)}{Nt}, \quad \mathscr{G}(t) = \log(1+t)+\log((1+t)^N-1)-\log(Nt),$$

* With the partial support of the Office of Naval Research and of the National Institutes of Health.

and

$$\mu_r = \frac{2}{N(r+1)} B_{r+1}^{(1)}\left(\frac{N+1}{2}\right) \quad (r \text{ even}), \qquad \kappa_r = (-1)^r \frac{B_r^{(1)}}{r}[N^r - 1] \quad (r \geqslant 2),$$

$$= 0 \quad (r \text{ odd}), \qquad \kappa_1 = \tfrac{1}{2}(N+1) = (-1)\, B_1^{(1)}[N+1],$$

$$\mu_{[r]} = \left(\frac{N+1}{r+1}\right)(N-1)^{(r-1)}, \qquad \kappa_{[r]} = \frac{1}{r}\left[(-1)^{r-1} r! + \sum_{j=0}^{r} {}^{r}C_j (-1)^j B_{r-j}^{(r)} B_j^{(1)} N^j\right].$$

These results are well known. The factorial cumulants do not appear to be capable of simpler expression.

THE POLYKAYS OF THE FIRST N NATURAL NUMBERS

The polykays $K_{rst\ldots}$ are functions of the moments (or cumulants) of the finite population composed of N elements. The same function, $k_{rst\ldots}$ of the sample of n averaged over all possible samples will be $K_{rst\ldots}$, i.e.

$$\mathscr{E}_N(k_{rst\ldots}) = K_{rst\ldots}.$$

For example, if μ_r and κ_r are the rth moment and cumulant quoted in the preceding section then

$$K_2 = \frac{N}{N-1}\mu_2 = \frac{N}{N-1}\kappa_2 = \frac{N(N+1)}{12},$$

or again

$$K_4 = \frac{1}{N^{(4)}}[N^3(N+1)\mu_4 - 3N^2(N-1)\mu_2^2] = -\frac{N^2(N+1)^2}{120}$$

and so on. The fact that the K's are calculated instead of the moments or cumulants is solely for ease of manipulation.

SINGLE SUFFIX POLYKAYS

From Aty's (1954) general result (a proof of which we have given elsewhere), for the rth central moment, $M(1^r)$, of the mean of a sample of n from a general finite population of N we see that $n^{r-1}M(1^r)$ is a polynomial in n, N and the polykays of the finite population. More precisely, considering this purely as a polynomial in n with constant coefficients we have that $n^{r-1}M(1^r) = K_r + \text{terms in } n^1 \text{ and higher powers}$. Further from the general relationship for cumulants in terms of moments we have that

$$\kappa(1^r) - M(1^r)$$

is a finite sum of products of at least two of the $\{M(1^s)\}$, the sum of the weights in each product being equal to r. Hence it is true generally that

$$K_r = \lim_{n \to 0} n^{r-1} K(1^r).$$

Since we know explicitly the rth cumulant of a sum, ω, of n elements chosen randomly from the finite population of N natural numbers we can find K_r for this population. The Eulerian identity, proved by Cauchy (1843), states in statistical terms that the p.g.f. of ω is

$$\Pi(z) = \frac{z^{\frac{1}{2}[n(n+1)]} \prod_{r=1}^{N} (1 - z^r)}{{}^{N}C_n \prod_{r=1}^{n} (1 - z^r) \prod_{r=1}^{N-n} (1 - z^r)}.$$

Accordingly since the rth cumulant of ω is

$$\kappa_r = 0 \quad r \geqslant 3, \quad (r \text{ odd}),$$

$$\kappa_r = \frac{1}{r(r+1)} B_r^{(1)} \{B_{r+1}^{(1)}(N+1) - B_{r+1}^{(1)}(N-n+1) - B_{r+1}^{(1)}(n+1) + B_{r+1}^{(1)}(1)\} \quad (r \text{ even}),$$

for even r

$$\lim_{n\to 0} n^{-1}\kappa_r = \frac{1}{r(r+1)} B_r^{(1)} \left[\frac{\partial}{\partial x} B_{r+1}^{(1)}(x+1)\right]_0^N,$$

i.e.

$$K_r = \frac{B_r^{(1)}}{r}[B_r^{(1)}(N+1) - B_r^{(1)}],$$

for odd r

$$K_r = 0, \quad r \geqslant 3; \qquad K_1 = \tfrac{1}{2}(N+1).$$

Unit suffix polykays

The polykays with suffices which have unity only as a part have a similar simple expression. It is known that

$$K_{1^r} = \frac{[1^r]}{N^{(r)}} = \frac{U_r}{N^{(r)}},$$

where U_r is the unitary symmetric function of weight r. We note that U_r is equal to

$$U_r = (-1)^r N^{(r)} B_r^{(N+1)}$$

and consequently

$$K_{1^r} = (-1)^r B_r^{(N+1)}.$$

Relations between the polykays for any finite population

The polykays with more than one suffix and with one or more of these suffices greater than unity do not seem generally to be expressible simply in terms of the Bernoulli polynomials of any order. Aty (unpublished thesis) by inverting Wishart's equations obtained formulae for the general finite population in which the polykays are expressed as functions of the single suffix kays. Either using his results or by going back to Wishart's original formulae the manipulation, although elementary, is cumbersome. We looked therefore at a further possibility, that of expressing polykays in terms of others of the same weight. This has advantages since, for example, the polykays with suffices containing unity only as a part are usually easy to calculate for any finite population. Some of our results are set out below. They are quite general and may be used for any population.

$$K_{2^s 1^r} = \frac{N^{(s)}}{(r+s)^{(s)}} \sum_{i=0}^{s} (-1)^i {}^sC_i \left(\frac{r+2i}{r+i}\right) \frac{(r+i)!}{r!} \frac{(r+s)!}{(r+s+i)!} K_{1^{r+s+i}} K_{1^{s-i}},$$

$$K_{2^s} = {}^NC_s \left\{K_{1^s}^2 + 2\sum_{i=1}^{s} \frac{(-1)^i {}^sC_i}{{}^{s+i}C_i} K_{s+i} K_{s-i}\right\},$$

$$K_{2^r} K_2 = \frac{r}{N} K_{4 2^{r-1}} - \frac{r(r-1)}{N(N-1)} K_{3^2 2^{r-2}} + \frac{N+2r-1}{N-1} K_{2^{r+1}},$$

$$K_{3 2^{s-1} 1^r} = \frac{N}{s}\left\{K_{2^s 1^r} K_1 - \frac{r}{N} K_{2^{s+1} 1^{r-1}} - K_{2^s 1^{r+1}}\right\},$$

$$K_{3^2 2^{r-2}} = \frac{N-1}{r(r-1)}\left\{\left(\frac{N+2r-1}{N-1} K_{2^{r+1}} - K_{2^r} K_2\right) - r(K_{3 2^{r-1} 1} - K_{3 2^{r-1}} K_1)\right\},$$

$$K_{3^r}K_2 = \frac{r}{N}K_{53^{r-1}} - \frac{r(r-1)}{N(N-1)}K_{4^2 3^{r-2}} + \frac{(N+6r-1)}{N-1}K_{3^r 2},$$

$$K_{41^s} = N\left[(K_{31^s}K_1 - K_{31^{s+1}}) + \frac{s}{2}\left(K_{2^2 1^s} + \frac{s-1}{N}K_{2^2 1^{s-2}} - K_{2^2 1^{s-1}}K_1\right)\right],$$

$$K_{42^{r-1}} = \frac{N-1}{r}\left[K_{2^r}K_2 - \frac{(N+2r-1)}{N-1}K_{2^{r+1}}\right] + [K_{32^{r-1}}K_1 - K_{32^{r-1}1}],$$

$$K_r K_1 = \frac{K_{r+1}}{N} + K_{r,1},$$

$$K_{r11} = K_r K_1^2 - \frac{K_{r+2}}{N^2} - \frac{2K_{r+1,1}}{N} - \frac{K_{r,2}}{N},$$

$$K_r K_{11} = \frac{2K_{r+1,1}}{N} - \frac{2}{N^{(2)}}\sum_{i=1}^{[\frac{1}{2}r]} {}^rC_{r-i}K_{r-i+1,i+1} + K_{r11} \quad (r \text{ odd}),$$

$$K_r K_{11} = \frac{2K_{r+1,1}}{N} - \frac{2}{N^{(2)}}\sum_{i=1}^{\frac{1}{2}r-1} {}^rC_{r-i}K_{r-i+1,i+1} - \frac{1}{N^{(2)}}{}^rC_{\frac{1}{2}r}K_{\frac{1}{2}r,\frac{1}{2}r} + K_{r11} \quad (r \text{ even}).$$

From the known expressions for the generalized Bernoulli polynomials (see, for example, Davis, 1935), the formulae of the previous sections may be used to obtain the polykays descriptive of the finite population of the first N natural numbers. These up to and including the eighth-order polykays are given in Table 1.

Table 1. *The polykays of the first N natural numbers*

Poly-kay	Dividing factor	Coefficients of N^8	N^7	N^6	N^5	N^4	N^3	N^2	N^1	N^0
K_1	2	—	—	—	—	—	—	—	1	1
K_2	12	—	—	—	—	—	—	1	1	—
K_{11}	12	—	—	—	—	—	—	3	5	2
K_3	—	—	—	—	—	—	0	0	0	0
K_{21}	24	—	—	—	—	—	1	2	1	—
K_{1^3}	8	—	—	—	—	—	1	2	1	—
K_4	−120	—	—	—	—	1	2	1	—	—
K_{31}	120	—	—	—	—	—	1	2	1	—
K_{22}	720	—	—	—	—	5	6	−5	−6	—
K_{21^2}	720	—	—	—	—	15	40	33	8	—
K_{1^4}	240	—	—	—	—	15	30	5	−18	−8
K_5	—	—	—	—	0	0	0	0	0	0
K_{41}	−240	—	—	—	1	3	3	1	—	—
K_{32}	—	—	—	—	0	0	0	0	0	0
K_{31^2}	120	—	—	—	—	1	3	3	1	—
$K_{2^2 1}$	1,440	—	—	—	15	30	5	−18	−8	—
K_{21^3}	480	—	—	—	5	15	13	1	−2	—
K_{1^5}	96	—	—	—	3	5	−5	−13	−6	—
K_6	504	—	—	2	6	5	0	−1	—	—
K_{51}	−504	—	—	—	2	6	5	0	−1	—
K_{42}	−10,080	—	—	7	5	−27	−25	20	20	—
K_{41^2}	−10,080	—	—	21	77	81	−9	−54	−20	—
K_{3^2}	−5,040	—	—	—	6	15	4	−15	−10	—
K_{321}	10,080	—	—	—	7	17	3	−17	−10	—
K_{2^3}	60,480	—	—	35	21	−157	−141	122	120	—
K_{31^3}	10,080	—	—	—	63	231	299	153	22	—
$K_{2^2 1^2}$	60,480	—	—	105	301	147	−257	−252	−44	—
K_{21^4}	20,160	—	—	105	315	161	−275	−322	−96	—
K_{1^6}	4,032	—	—	63	63	−315	−539	−74	236	96

Table 1 (*cont.*)

Polykay	Dividing factor	Coefficients of N^8	N^7	N^6	N^5	N^4	N^3	N^2	N^1	N^0
K_7	—	—	0	0	0	0	0	0	0	0
K_{61}	1,008	—	2	8	11	5	−1	−1	—	—
K_{52}	—	—	0	0	0	0	0	0	0	0
K_{43}	—	—	0	0	0	0	0	0	0	0
K_{51^2}	−504	—	—	2	8	11	5	−1	−1	—
K_{421}	−20,160	—	7	12	−22	−52	−5	40	20	—
K_{32^2}	—	—	0	0	0	0	0	0	0	0
$K_{3^2 1}$	−10,080	—	—	6	21	19	−11	−25	−10	—
K_{41^3}	−6,720	—	7	28	18	−68	−133	−88	−20	—
K_{321^2}	10,080	—	—	7	24	20	−14	−27	−10	—
$K_{2^3 1}$	120,960	—	35	56	−136	−298	−19	242	120	—
K_{31^4}	5,040	—	—	21	84	110	32	−35	−20	—
$K_{2^2 1^3}$	40,320	—	35	112	56	−138	−131	26	40	—
K_{21^5}	8,064	—	21	56	−28	−210	−201	−38	16	—
K_{1^7}	1,152	—	9	0	−84	−98	91	194	80	—
K_8	−720	3	12	14	0	−7	0	2	—	—
K_{71}	720	—	3	12	14	0	−7	0	2	—
K_{62}	30,240	10	4	−71	−53	145	133	−84	−84	—
K_{53}	5,040	—	3	7	−7	−28	−10	21	41	—
K_{4^2}	100,800	7	−20	−62	180	503	120	−448	−280	—
K_{61^2}	30,240	30	140	155	−183	−475	−175	146	84	—
K_{521}	−30,240	—	10	22	−29	−95	−23	73	42	—
K_{431}	−100,800	—	7	40	78	40	−57	−80	−28	—
K_{42^2}	−604,800	35	−48	−380	42	1,225	678	−880	−672	—
$K_{3^2 2}$	−302,400	—	30	63	−103	−325	−95	262	168	—
K_{51^3}	−10,080	—	30	63	−103	−325	−95	262	168	—
K_{421^2}	−604,800	105	250	−344	−1,410	−785	960	1,024	200	—
$K_{3^2 1^2}$	−302,400	—	90	333	297	−245	−455	−88	68	—
$K_{32^2 1}$	604,800	—	35	72	−128	−370	−75	298	168	—
K_{2^4}	3,628,800	175	−140	−1,754	−176	−5,335	3,340	−3,756	−3,024	—
K_{321^3}	604,800	—	35	1,290	1,778	−380	−1,765	−910	−48	—
K_{41^4}	−201,600	105	420	−160	2,256	6,145	4,830	1,480	64	—
$K_{2^3 1^2}$	3,628,800	525	1,190	−1,690	−6,052	−2,275	4,670	3,440	192	—
K_{31^5}	120,960	—	315	1,260	1,200	−1,418	−3,419	−2,146	−400	—
$K_{2^2 1^4}$	241,920	105	336	14	− 940	−795	444	676	160	—
K_{21^6}	241,920	315	630	−1,890	−6,160	−3,865	3,226	4,288	1,152	—
K_{1^8}	34,560	135	−180	−1,890	−840	6,055	8,140	884	−3,088	−1,152

The polykays of the hypergeometric distribution

Suppose a finite population of β black balls and $N-\beta$ white balls all indistinguishable except as regards colour. We may regard this as a finite population of β 1's and $(N-\beta)$ 0's for sampling purposes. Since the result $K_r = \lim_{n\to 0} n^{r-1} M(1^r)$ is quite general, application in this context yields

$$K_r = \sum_{j=1}^{r} \frac{\Delta^j 0^r}{j!} (-1)^{j-1} \frac{\beta^{(j)}}{N^{(j)}}$$

and in particular

$$K_1 = \frac{\beta}{N}, \quad K_2 = \frac{\beta(N-\beta)}{N^{(2)}}, \quad K_3 = \frac{\beta(N-\beta)(N-2\beta)}{N^{(3)}},$$

$$K_4 = \frac{\beta(N-\beta)}{N^{(4)}} [N(N+1) - 6\beta(N-\beta)].$$

The multiple suffix polykay of unit parts is

$$K_{1^r} = \frac{\beta^{(r)}}{N^{(r)}}.$$

From these quantities all the intervening polykays may be deduced from our formulae and in particular

$$K_{22} = \frac{\beta^{(2)}(N-\beta)^{(2)}}{N^{(4)}}.$$

Using Aty's result for the mean of a sample it is possible to write down the moments of the distribution of the number, a, of black balls in a sample of m drawn from N without replacement.

MOOD'S CRITERION

For the sake of example and to illustrate the use of the formulae which we have provided we consider the moments of a criterion proposed by Mood (1954). Given a random sequence of N elements, n of one kind and $N-n$ of another, which are ranked in order of magnitude, Mood proposed as a possible test for difference of dispersion the average of the sum of the squares of the deviations of the ranks of the elements of one kind from $\frac{1}{2}(N+1)$. In general terms we may write this as

$$\theta = \frac{1}{n}\sum_{i=1}^{n}(x_i - K_1)^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - k_1 + k_1 - K_1)^2,$$

where k_1 is the sample mean and K_1 the population mean of the N elements. Hence for any finite population, exclusive of any idea of ranking,

$$\theta = \frac{n-1}{n}k_2 + (k_1 - K_1)^2,$$

or

$$\theta - K_1^2 = k_2 + k_{11} - 2k_1K_1 = \Lambda \quad \text{(say)}.$$

The moments of Λ are straightforward. Thus

$$\mathscr{E}(\Lambda) = K_2 + K_{11} - 2K_1^2,$$

$$\mathscr{E}(\Lambda - \mathscr{E}(\Lambda))^2 = \left(\frac{1}{n} - \frac{1}{N}\right)\left[K_4\left(\frac{N-2}{N}\right)^2 + 2K_{22}\left(\frac{N-2}{N}\right)\right],$$

$$\mathscr{E}(\Lambda - \mathscr{E}(\Lambda))^3 = \left(\frac{1}{n} - \frac{1}{N}\right)\left(\frac{1}{n} - \frac{2}{N}\right) \times \left[K_3\left(\frac{N-2}{N}\right)^3 + 12K_{42}\frac{(N-2)^{(2)}}{N^2} + K_{33}\frac{(10N^2 - 48N + 64)}{N^2} + 8K_{2^3}\frac{(N-3)}{N}\right]$$

with a longer expression for the fourth moment which we do not reproduce here. For the particular case considered by Mood, the finite population is the first N integers, $K_1 = \frac{1}{2}(N+1)$ and the polykays are those which we have tabulated. On substitution we have,

$$\mathscr{E}(\theta) = \frac{N^2-1}{12},$$

$$\mathscr{E}(\theta - \mathscr{E}(\theta))^2 = \frac{N(N+1)(N^2-4)}{180}\left(\frac{1}{n} - \frac{1}{N}\right),$$

$$\mathscr{E}(\theta-\mathscr{E}(\theta))^3 = \frac{N^2(N+1)(N+2)(N^2-16)}{3780}\left(\frac{1}{n}-\frac{1}{N}\right)\left(\frac{1}{n}-\frac{2}{N}\right),$$

$$\mathscr{E}(\theta-\mathscr{E}(\theta))^4 = \frac{N(N+1)(N+2)}{75\,600}\left\{2N(N+1)(N^3-4N^2+20N+88)\left(\frac{1}{n^3}-\frac{4}{n^2N}+\frac{6}{nN^2}-\frac{3}{N^3}\right)\right.$$
$$\left.+(N-1)(7N^4-20N^3-4N^2-64N-528)\left(\frac{1}{n}-\frac{1}{N}\right)^2\right\}.$$

Alternative approach

An alternative method of approach which is suitable for some rank criteria and which is applicable to finding the moments of Mood's statistic is to make use of Aty's general results. Aty (1954) showed for $r \leqslant 12$ that for any finite population of N elements which has population polykays $K_{rs\ldots}$ the moments of the mean k_1 of a sample of n

$$M(1^r) = \mathscr{E}_N(k_1-K_1)^r$$

may be written as

$$M(1^2) = K_2A_2, \quad M(1^3) = K_3A_3,$$
$$M(1^4) = K_4A_4+3K_{22}A_2^2$$

and so on, where

$$A_r = \alpha^{r-1}-\frac{\alpha^{r-2}}{N}+\frac{\alpha^{r-3}}{N^2}\ldots(-1)^{r-2}\frac{\alpha}{N^{r-2}}$$

and

$$\alpha = \frac{1}{n}-\frac{1}{N}.$$

Mood's statistic can be regarded as a mean of n elements each randomly drawn without replacement from a population

$$(-r)^2,\ (-r+1)^2,\ \ldots,\ (-1)^2,\ 0,\ (1)^2,\ \ldots,\ (r)^2 \quad (N = 2r+1),$$

or

$$(-r+\tfrac{1}{2})^2,\ (-r+\tfrac{3}{2})^2,\ \ldots,\ (-\tfrac{1}{2})^2,\ (\tfrac{1}{2})^2,\ \ldots,\ (r-\tfrac{1}{2})^2 \quad (N = 2r).$$

The polykays of these two populations are the same functions of N. They are

$$K_1 = \frac{N^2-1}{12}, \quad K_2 = \frac{N(N+1)(N^2-4)}{180},$$

$$K_3 = \frac{N^2(N+1)(N+2)(N^2-16)}{3780}, \quad K_4 = \frac{N^2(N+1)^2(N+2)(N^3-4N^2+20N+88)}{37{,}800},$$

$$K_{22} = \frac{N(N+2)(N^2-1)}{226{,}800}(7N^4-20N^3-4N^2-64N-528).$$

Consequently from Aty's results the moments of Mood's statistic follow. It will be noticed that if $n = \lambda N$, where λ is constant, the limiting value of the moment ratio β_2 is 3.

References

Abdel-Aty, S. H. (1954). *Biometrika*, **41**, 253–60.
Cauchy, A. (1843). *Comptes Rendus*, **17**, 523–31.
David, F. N. (1956). *Biometrika*, **43**, 485–8.
Davis, H. T. (1935). *Tables of the Higher Mathematical Functions*, **2**. Indiana: Principia Press. Inc.
Mood, A. M. (1954). *Ann. Math. Statist.* **25**, 514–22.
Nörlund, N. E. (1920). *Acta Math.* **43**, 121–96.
Wishart, J. (1952). *Biometrika*, **39**, 1–13.

Biometrika (1960), **47**, 1 *and* 2, *p.* 61
Printed in Great Britain

The distribution of regression coefficients in samples from bivariate non-normal populations

I. Theoretical investigation

By A. B. L. SRIVASTAVA

Statistical Laboratory, Indian Institute of Technology, Kharagpur

1. Introduction

The work done so far on the sampling distributions of statistics arising from bivariate non-normal populations mainly relates to the study of the effect of non-normality on r, the correlation coefficient. Pearson (1929) showed that considerable non-normality can be allowed in the parent bivariate population without seriously affecting the distribution of r; Quensel (1938) derived the density function of r in samples from a bivariate Gram–Charlier Type A population under the assumption that the population correlation ρ is small, and Gayen (1951) studied the distribution r in samples from populations represented by the Edgeworth surface for the case when ρ is not necessarily small. So far as the regression coefficient $b_{21} = m_{11}/m_{20}$ is concerned, only expressions for its first four cumulants have been obtained in the non-normal case. Quensel (1938) obtained the expressions for the mean and the variance of b_{21} from the joint characteristic function of the second-order moments and Cook (1951) derived the first four moments of b_{21}, using some generalizations of Kendall's k-statistics. It appears that no attempt has been made to study the frequency distribution of b_{21} when the parent population is non-normal, except for the work of Hyrenius (1952) who has obtained the distribution of b_{21} by considering the parent population to be specified by a bivariate compound normal distribution. Hill (1954) has derived cumulants of the function $m_{11} - cm_{20}$ using the k-statistics in the non-normal case, and has made use of them to calculate approximate percentage points of b_{21}.

Assuming the parent population to be specified by the Edgeworth surface (including terms up to those in $\lambda_{30}^2, \lambda_{30}\lambda_{21}, \ldots, \lambda_{03}^2$), the distribution of b_{21} has been derived here. It holds for the values of ρ not necessarily small, as well as (i) for any size of sample provided that fifth- and higher-order population cumulants are negligible, and (ii) for any population if samples are sufficiently large. The distribution of the t-statistic used for testing the significance of the regression coefficient has also been derived, and of the two cases of known and unknown marginal variances, the latter has been considered in detail because of its obvious practical importance. For this, in addition to the normal theory tail probability of t, the corrective terms in $\lambda_{40}, \lambda_{31}, \ldots, \lambda_{04}$ and $\lambda_{30}^2, \lambda_{30}\lambda_{21}, \ldots, \lambda_{03}^2$ have been obtained explicitly.

2. Frequency function of the regression coefficient

Let us consider the bivariate parent population to be represented by the Edgeworth surface

$$F(x,y) = \left\{1 + \sum_{i+j=3,4,6}\sum A_{ij}\frac{(-1)^{i+j}}{i!\,j!}\frac{\partial^{i+j}}{\partial x^i \partial y^j}\right\}\phi(x,y), \tag{1}$$

where

$$\phi(x,y) = \frac{1}{2\pi\sqrt{(\kappa_{20}\kappa_{02} - \kappa_{11}^2)}}\exp\left[-\frac{1}{2}\left(\frac{\kappa_{02}x^2 - 2\kappa_{11}xy + \kappa_{20}y^2}{\kappa_{20}\kappa_{02} - \kappa_{11}^2}\right)\right], \tag{2}$$

$$\left.\begin{aligned} A_{30} = \kappa_{30}, \quad A_{21} = \kappa_{21}, \quad A_{12} = \kappa_{12}, \\ A_{40} = \kappa_{40}, \quad A_{31} = \kappa_{31}, \quad A_{22} = \kappa_{22}, \\ A_{60} = 10\kappa_{30}^2, \quad A_{51} = 10\kappa_{30}\kappa_{21}, \quad A_{42} = 4\kappa_{30}\kappa_{12} + 6\kappa_{21}^2, \quad A_{33} = \kappa_{30}\kappa_{03} + 9\kappa_{21}\kappa_{12}, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \end{aligned}\right\} \tag{3}$$

and the κ_{ij}'s denote the semi-invariants of the population. We shall use the standardized semi-invariants $\lambda_{ij} = \kappa_{ij}/(\kappa_{20}^{\frac{1}{2}i}\kappa_{02}^{\frac{1}{2}j})$ in the following derivations and denote the population correlation coefficient $\kappa_{11}/\sqrt{(\kappa_{20}\kappa_{02})}$ by ρ.

Let (x_i, y_i), $(i = 1, 2, \ldots, N)$, be a random sample from the population represented by (1). Using Gayen's (1951) formula (9) and taking λ_{ij}'s to be the standardized semi-invariants, the joint characteristic function of

$$m_{20} = \frac{\Sigma(x_i - \bar{x})^2}{N}, \quad m_{11} = \frac{\Sigma(x_i - \bar{x})(y_i - \bar{y})}{N} \quad \text{and} \quad m_{02} = \frac{\Sigma(y_i - \bar{y})^2}{N}, \tag{4}$$

where

$$\bar{x} = \frac{\Sigma x_i}{N}, \quad \bar{y} = \frac{\Sigma y_i}{N},$$

can be obtained by substituting in it the following expressions for his D, τ_{20}, τ_{11} and τ_{02}

$$\left.\begin{aligned} D &= \left(1 - \frac{2}{N}\kappa_{20} it_{20}\right)\left(1 - \frac{2}{N}\kappa_{02} it_{02}\right) - \frac{4\kappa_{11}^2}{N}(it_{20})(it_{02}) - \frac{2\kappa_{11}}{N}(it_{11}) - \frac{(\kappa_{20}\kappa_{02} - \kappa_{11}^2)}{N^2}(it_{11})^2, \\ \tau_{20} &= \frac{1}{D} - \frac{2}{ND}(\rho^2\kappa_{20} it_{20} + \kappa_{11} it_{11} + \kappa_{02} it_{02}) - 1, \\ \tau_{11} &= \rho - \frac{\rho}{D} + \frac{1}{ND}(2\rho\kappa_{20} it_{20} + (1+\rho^2)\sqrt{(\kappa_{20}\kappa_{02})}\, it_{11} + 2\rho\kappa_{02} it_{02}), \\ \tau_{02} &= \frac{1}{D} - \frac{2}{ND}(\kappa_{20} it_{20} + \kappa_{11} it_{11} + \rho^2\kappa_{02} it_{02}) - 1. \end{aligned}\right\} \tag{5}$$

Gayen (1951, pp. 221–3) derived the joint distribution of m_{20}, m_{11} and m_{02} from the characteristic function expressed by his formula (9) relating to the case when $\kappa_{20} = \kappa_{02} = 1$. But in deriving the distribution of the regression coefficient $b_{21} = m_{11}/m_{20}$ it seems desirable not to impose any restrictions on κ_{20} and κ_{02}. However, when $\kappa_{20} \neq \kappa_{02}$, the joint frequency function of m_{20}, m_{11} and m_{02} can be derived without difficulty from Gayen's characteristic function (9) modified by our equations (5), and can be expressed by his formula (21)* with the following modifications:

(i) the normal theory term $f_0(m_{20}, m_{11}, m_{02})$ occurring in it becomes

$$f_0(m_{20}, m_{11}, m_{02}) = \frac{N^{N-1}}{4\pi\Gamma(N-2)(\kappa_{20}\kappa_{02} - \kappa_{11}^2)^{\frac{1}{2}(N-1)}}(m_{20}m_{02} - m_{11}^2)^{\frac{1}{2}(N-4)} \times \exp\left[-\tfrac{1}{2}N\left(\frac{\kappa_{02}m_{20} - 2\kappa_{11}m_{11} + \kappa_{20}m_{02}}{\kappa_{20}\kappa_{02} - \kappa_{11}^2}\right)\right], \tag{6}$$

(ii) the quantities M_{20}, M_{11} and M_{02} are now given by

$$\left.\begin{aligned} (\kappa_{20}\kappa_{02} - \kappa_{11}^2) M_{20} &= \kappa_{02}m_{20} - 2\kappa_{11}m_{11} + \rho^2\kappa_{20}m_{02}, \\ (\kappa_{20}\kappa_{02} - \kappa_{11}^2) M_{11} &= \rho\kappa_{02}m_{20} - (1+\rho^2)\sqrt{(\kappa_{20}\kappa_{02})}\, m_{11} + \rho\kappa_{20}m_{02}, \\ (\kappa_{20}\kappa_{02} - \kappa_{11}^2) M_{02} &= \rho^2\kappa_{02}m_{20} - 2\kappa_{11}m_{11} + \kappa_{20}m_{02}, \end{aligned}\right\} \tag{7}$$

* In Gayen's (1951, pp. 222–3) result (21), inside the square brackets which have $-6\lambda_{03}\lambda_{12}$ as the coefficient, M_{20} should read as M_{02}. Incidentally, it may be pointed out that in his formula (39) for var (r) on p. 226, in the second line, the term $\frac{1}{4}(1-\rho^2)(2+11\rho^2)$ is a misprint for $\frac{1}{4}(1-\rho^2)^2(2+11\rho^2)$.

and (iii) the quantity $m_{20}m_{02}-m_{11}^2$ occurring in the formula (21) is replaced by

$$(m_{20}m_{02}-m_{11}^2)/(\kappa_{20}\kappa_{02}).$$

Let us denote his result (21) so modified by

$$g(m_{20}, m_{11}, m_{02}). \tag{8}$$

In (8), by substituting for M_{20}, M_{11} and M_{02} the expressions given by (7), a typical term can be seen to be of the form

$$f_0(m_{20}, m_{11}, m_{02})\, m_{20}^{r_1} m_{11}^{r_2} m_{02}^{r_3}, \tag{9}$$

r_1, r_2, r_3 being positive integers such that $r = r_1+r_2+r_3 \leqslant 3$.

Putting $m_{20}b_{21} = m_{11}$ and integrating out m_{20} and m_{02}, we find the frequency function of b_{21} corresponding to (9) is given by

$$f_{r_1r_2r_3}(b_{21}) = f_0(b_{21})\left(\frac{2}{N}\right)^r (\kappa_{20}\kappa_{02}-\kappa_{11}^2)^r \times \sum_{j=0}^{r}\left[{}_{r_3}C_j\, b_{21}^{r_2+2j}\frac{\Gamma(\frac{1}{2}N-1+r_3-j)\,\Gamma(\frac{1}{2}N+r_1+r_2+j)}{\Gamma(\frac{1}{2}N)\,\Gamma(\frac{1}{2}N-1)\,\kappa_{20}^{r_3-j}(\kappa_{20}b_{21}^2-2\kappa_{11}b_{21}+\kappa_{02})^{r_1+r_2+j}}\right], \tag{10}$$

where

$$f_0(b_{21}) = \frac{2^{N-3}\Gamma(\frac{1}{2}N-1)\,\Gamma(\frac{1}{2}N)\,(\kappa_{20}\kappa_{02}-\kappa_{11}^2)^{\frac{1}{2}(N-1)}}{\pi\Gamma(N-2)\,\kappa_{20}^{\frac{1}{2}N-1}(\kappa_{20}b_{21}^2-2\kappa_{11}b_{21}+\kappa_{02})^{\frac{1}{2}N}}, \tag{11}$$

is the normal theory frequency function of b_{21}.

Using (10), we obtain from (8) after some simplification, the frequency function of b_{21} as

$$\begin{aligned}
f(b_{21}) = f_0(b_{21})\Bigg[1 &+ \frac{N-1}{8N(N+1)(1-\rho^2)^2}\Bigg\{\frac{N(N+2)}{B_{11}^2}\\
&\times(\lambda_{40}B_{10}^4-4\lambda_{31}B_{10}^3B_{01}+6\lambda_{22}B_{10}^2B_{01}^2-4\lambda_{13}B_{10}B_{01}^3+\lambda_{04}B_{01}^4)\\
&-\frac{2N}{B_{11}}[\{N(1-\rho^2)+2\rho^2+1\}(\lambda_{40}B_{10}^2-2\lambda_{31}B_{10}B_{01}+\lambda_{22}B_{01}^2)\\
&-6\rho(\lambda_{31}B_{10}^2-2\lambda_{22}B_{10}B_{01}+\lambda_{13}B_{01}^2)+3(\lambda_{22}B_{10}^2-2\lambda_{13}B_{10}B_{01}+\lambda_{04}B_{01}^2)]\\
&+[\{N^2(1-\rho^2)^2+2N(1-\rho^2)\rho^2+4\rho^2-1\}\lambda_{40}-4\rho\{N(1-\rho^2)+2\rho^2+1\}\lambda_{31}\\
&+2\{N(1-\rho^2)+8\rho^2+1\}\lambda_{22}-12\rho\lambda_{13}+3\lambda_{04}]\Bigg\}\\
&+\frac{N-2}{12N(N+1)(N+3)(1-\rho^2)^3}\Bigg\{\frac{N(N+2)(N+4)}{B_{11}^3}[\lambda_{30}^2B_{10}^6-6\lambda_{30}\lambda_{21}B_{10}^5B_{01}\\
&+3(2\lambda_{30}\lambda_{12}+3\lambda_{21}^2)B_{10}^4B_{01}^2-2(\lambda_{30}\lambda_{03}+9\lambda_{21}\lambda_{12})B_{10}^3B_{01}^3\\
&+3(3\lambda_{12}^2+2\lambda_{03}\lambda_{21})B_{10}^2B_{01}^4-6\lambda_{03}\lambda_{12}B_{10}B_{01}^5+\lambda_{03}^2B_{01}^6]\\
&-\frac{3N(N+2)}{B_{11}^2}[\{N(1-\rho^2)+2\rho^2+3\}\{\lambda_{30}^2B_{10}^4-4\lambda_{30}\lambda_{21}B_{10}^3B_{01}\\
&+2(\lambda_{30}\lambda_{12}+2\lambda_{21}^2)B_{10}^2B_{01}^2-4\lambda_{21}\lambda_{12}B_{10}B_{01}^3+\lambda_{12}^2B_{01}^4\}\\
&-10\rho\{\lambda_{30}\lambda_{21}B_{10}^4-2(\lambda_{30}\lambda_{12}+\lambda_{21}^2)B_{10}^3B_{01}+(\lambda_{30}\lambda_{03}+5\lambda_{21}\lambda_{12})B_{10}^2B_{01}^2\\
&-2(\lambda_{12}^2+\lambda_{03}\lambda_{21})B_{10}B_{01}^3+\lambda_{03}\lambda_{12}B_{01}^4\}+5\{\lambda_{21}^2B_{10}^4-4\lambda_{21}\lambda_{12}B_{10}^3B_{01}\\
&+2(2\lambda_{12}^2+\lambda_{03}\lambda_{21})B_{10}^2B_{01}^2-4\lambda_{03}\lambda_{12}B_{10}B_{01}^3+\lambda_{03}^2B_{01}^4\}\\
&-2(1-\rho^2)^2\kappa_{02}\{(\lambda_{21}^2-\lambda_{30}\lambda_{12})B_{10}^2+(\lambda_{30}\lambda_{03}-\lambda_{21}\lambda_{12})B_{10}B_{01}+(\lambda_{12}^2-\lambda_{03}\lambda_{21})B_{01}^2\}]\\
&+\frac{3N}{B_{11}}[\{N(N-2)(1-\rho^2)^2+6(N-2)(1-\rho^2)+15\}(\lambda_{30}^2B_{10}^2-2\lambda_{30}\lambda_{21}B_{10}B_{01}+\lambda_{21}^2B_{01}^2)\\
&-12\rho\{N(1-\rho^2)+2\rho^2+3\}\{\lambda_{30}\lambda_{21}B_{10}^2-(\lambda_{30}\lambda_{12}+\lambda_{21}^2)B_{10}B_{01}+\lambda_{21}\lambda_{12}B_{01}^2\}
\end{aligned}$$

$$+30\rho^2\{\lambda_{30}\lambda_{12}B_{10}^2-(\lambda_{30}\lambda_{03}+\lambda_{21}\lambda_{12})B_{10}B_{01}+\lambda_{03}\lambda_{21}B_{01}^2\}$$
$$+6\{N(1-\rho^2)+7\rho^2+3\}(\lambda_{21}^2B_{10}^2-2\lambda_{21}\lambda_{12}B_{10}B_{01}+\lambda_{12}^2B_{01}^2)$$
$$-60\rho\{\lambda_{21}\lambda_{12}B_{10}^2-(\lambda_{12}^2+\lambda_{03}\lambda_{21})B_{10}B_{01}+\lambda_{03}\lambda_{12}B_{01}^2\}$$
$$+15(\lambda_{12}^2B_{10}^2-2\lambda_{03}\lambda_{12}B_{10}B_{01}+\lambda_{03}^2B_{01}^2)$$
$$-2(1-\rho^2)^2\kappa_{02}\{(N(1-\rho^2)+3)(\lambda_{21}^2-\lambda_{30}\lambda_{12})+3\rho(\lambda_{30}\lambda_{03}-\lambda_{21}\lambda_{12})+3(\lambda_{12}^2-\lambda_{03}\lambda_{21})\}]$$
$$-[\{N(N^2-4)(1-\rho^2)^3+3N(N-2)(1-\rho^2)^2+9(N-2)(1-\rho^2)+15\}\lambda_{30}^2$$
$$-6\rho\{N(N-2)(1-\rho^2)^2+6(N-2)(1-\rho^2)+15\}\lambda_{30}\lambda_{21}$$
$$+18\rho^2\{N(1-\rho^2)+2\rho^2+3\}\lambda_{30}\lambda_{12}+3\{(N-2)(N-6)(1-\rho^2)^2$$
$$+6(2N-9)(1-\rho^2)+45\}\lambda_{21}^2-30\rho^3\lambda_{30}\lambda_{03}-18\rho\{2N(1-\rho^2)+9\rho^2+6\}\lambda_{21}\lambda_{12}$$
$$+9\{N(1-\rho^2)+12\rho^2+3\}\lambda_{12}^2+90\rho^2\lambda_{03}\lambda_{21}-90\rho\lambda_{03}\lambda_{12}+15\lambda_{03}^2]\Big\}\Big], \tag{12}$$

where

$$\left.\begin{aligned} B_{10} &= \rho\sqrt{\kappa_{20}}b_{21}-\sqrt{\kappa_{02}},\\ B_{01} &= \sqrt{\kappa_{20}}b_{21}-\rho\sqrt{\kappa_{02}},\\ B_{11} &= \kappa_{20}b_{21}^2-2\kappa_{11}b_{21}+\kappa_{02}. \end{aligned}\right\} \tag{13}$$

As a check, it may be shown that the integral of (12) over $(-\infty,\infty)$ is unity. The corrective terms due to the various λ's do not appear to be expressible as derivatives with respect to $\beta_{21}\,(=\kappa_{11}/\kappa_{20})$ or ρ, of the normal theory frequency function of b_{21}, and in view of this further simplification of the frequency function of b_{21} as in the case of r (Gayen, 1951) has not been attempted.

3. Mean and variance of the regression coefficient

We note that the typical term of (12) is of the form

$$C\frac{(\rho\sqrt{\kappa_{20}}b_{21}-\sqrt{\kappa_{02}})^{r_1}(\sqrt{\kappa_{20}}b_{21}-\rho\sqrt{\kappa_{02}})^{r_2}}{(\kappa_{20}b_{21}^2-2\kappa_{11}b_{21}+\kappa_{02})^s}f_0(b_{21}), \tag{14}$$

r_1, r_2 taking the values $0, 1, 2, \ldots, 6$ such that $r_1+r_2\leqslant 6$ and s taking the values $0, 1, 2, 3$.

It can be shown that kth moment about zero of b_{21} corresponding to (14) is given by

$$\mu_k'(b_{21})_{r_1r_2s}=C\frac{\kappa_{02}^{\frac{1}{2}(k+r_1+r_2-2s)}(1-\rho^2)^{\frac{1}{2}(r_1+r_2-2s)}}{\kappa_{20}^{\frac{1}{2}k}B(\frac{1}{2},\frac{1}{2}(N-1))}\int_{-\infty}^{\infty}\frac{[u\sqrt{(1-\rho^2)}+\rho]^k[\rho u-\sqrt{(1-\rho^2)}]^{r_1}u^{r_2}}{(1+u^2)^{\frac{1}{2}N+s}}du. \tag{15}$$

The evaluation of the integral in (15) is quite straightforward. With its help we can derive the mean and the variance of b_{21} directly from (12). They are found to be

$$E(b_{21})=\left(\frac{\kappa_{02}}{\kappa_{20}}\right)^{\frac{1}{2}}\left[\rho+\frac{N-1}{N(N+1)}(\rho\lambda_{40}-\lambda_{31})-\frac{4(N-2)}{N(N+1)(N+3)}(\rho\lambda_{30}^2-\lambda_{30}\lambda_{21})\right] \tag{16}$$

and

$$V(b_{21})=\frac{\kappa_{02}}{(N-3)\kappa_{20}}\Big[(1-\rho^2)+\frac{N-1}{N(N+1)}\{[(N-6)\rho^2+1]\lambda_{40}-(N-5)(2\rho\lambda_{31}-\lambda_{22})\}$$
$$-\frac{4(N-2)}{N(N+1)(N+3)}\{[2(N-5)\rho^2+1]\lambda_{30}^2-2(2N-9)\rho\lambda_{30}\lambda_{21}$$
$$+(N-6)\lambda_{21}^2+(N-3)\lambda_{30}\lambda_{12}\}\Big]. \tag{17}$$

We shall now compare the formulae (16) and (17) with the well-known results of Quensel (1938) and Cook (1951). Quensel derived the formulae for the mean and the variance of the regression coefficient from the joint characteristic function of the second-order moments which is valid for small ρ only. Though his final expressions include terms up to those of

order N^{-1} only, from his formulae (133) and (134) (Quensel, 1938, pp. 106–7) the full expressions for $E(b_{21})$ and $V(b_{21})$ can easily be written retaining, however, only those terms which involve cumulants (and products and squares of cumulants) relevant to the present work. Keeping in view the fact that in Quensel's work the coefficients of the order $\rho^2\lambda_{30}^2$ have not been taken into account, we find that these expressions agree with our formulae (16) and (17) except for the term in $\rho^3\lambda_{30}^2$ in (17). It is thus observed that the mean and the variance of b_{21}, check with those derived by Quensel.

If we express our results (16) and (17) in powers of $1/N$ and retain terms up to those of the order N^{-2}, we get

$$E(b_{21}) = \left(\frac{\kappa_{02}}{\kappa_{20}}\right)^{\frac{1}{2}}\left[\rho + \frac{N-2}{N^2}(\rho\lambda_{40} - \lambda_{31}) - \frac{4}{N^2}(\rho\lambda_{30}^2 - \lambda_{30}\lambda_{21})\right] \tag{18}$$

and

$$V(b_{21}) = \frac{\kappa_{02}}{\kappa_{20}}\left[\frac{N+3}{N^2}(1-\rho^2) + \frac{1}{N}(\rho^2\lambda_{40} - 2\rho\lambda_{31} + \lambda_{22})\right.$$
$$\left. - \frac{1}{N^2}\{(5\rho^2-1)\lambda_{40} - 8\rho\lambda_{31} + 4\lambda_{22}\} - \frac{4}{N^2}(2\rho^2\lambda_{30}^2 - 4\rho\lambda_{30}\lambda_{21} + \lambda_{30}\lambda_{12} + \lambda_{21}^2)\right]. \tag{19}$$

Cook's formulae for the mean and the variance of b_{21}, up to the order N^{-2} for a general bivariate population with cumulants existing derived by using the k-statistics, agree with our results (18) and (19), provided only those semi-invariants (and their products and squares) which have been considered in the present study, are retained in them.*

We may now consider a few particular cases of the population form. If the variate x has a normal distribution (so that $\lambda_{i0} = 0$), we have

$$E(b_{21}) = \left(\frac{\kappa_{02}}{\kappa_{20}}\right)^{\frac{1}{2}}\left[\rho - \frac{N-1}{N(N+1)}\lambda_{31}\right], \tag{20}$$

and

$$V(b_{21}) = \frac{\kappa_{02}}{(N-3)\kappa_{20}}\left[(1-\rho^2) - \frac{(N-1)(N-5)}{N(N+1)}(2\rho\lambda_{31} - \lambda_{22}) - \frac{4(N-2)(N-6)}{N(N+1)(N+3)}\lambda_{21}^2\right]. \tag{21}$$

In the case of strictly linear regression, when $\lambda_{i1} = \rho\lambda_{i+1,0}$ (Wicksell, 1933), the formulae (16) and (17) reduce to

$$E(b_{21}) = \left(\frac{\kappa_{02}}{\kappa_{20}}\right)^{\frac{1}{2}}\rho = \frac{\kappa_{11}}{\kappa_{20}}, \tag{22}$$

and

$$V(b_{21}) = \frac{\kappa_{02}}{(N-3)\kappa_{20}}\left[(1-\rho^2) - \frac{N-1}{N(N+1)}\{[(N-4)\rho^2 - 1]\lambda_{40} - (N-5)\lambda_{22}\}\right.$$
$$\left. + \frac{4(N-2)}{N(N+1)(N+3)}\{[(N-2)\rho^2 - 1]\lambda_{30}^2 - (N-3)\lambda_{30}\lambda_{12}\}\right]. \tag{23}$$

If the two variates x and y are uncorrelated, then $\lambda_{ij} = 0$ for $i \neq 0, j \neq 0$, and consequently

$$E(b_{21}) = 0 \tag{24}$$

and

$$V(b_{21}) = \frac{\kappa_{02}}{(N-3)\kappa_{20}}\left[1 + \frac{N-1}{N(N+1)}\lambda_{40} - \frac{4(N-2)}{N(N+1)(N+3)}\lambda_{30}^2\right]. \tag{25}$$

* It may be remarked here that since Cook's formulae for the mean and the variance of b_{21} contain all the terms of the order N^{-1} and N^{-2} (e.g. those in λ_{60}, λ_{51}, λ_{40}^2, $\lambda_{40}\lambda_{31}$, etc.), they will be more accurate where our assumptions about the parent population are not true and where in addition to the knowledge on λ_{30}, λ_{40}, etc., we have some knowledge on λ_{60}, λ_{51}, etc., also available, provided N is not very small.

If, in addition, x is a normal variate

$$V(b_{21}) = \frac{\kappa_{02}}{(N-3)\kappa_{20}}, \tag{26}$$

which is the same as the variance of b_{21} when both the variates are independent and normally distributed.

Quensel discussed some of these special cases, retaining terms up to order N^{-1} only in his formulae. We find them to check with the above results.

4. TEST OF SIGNIFICANCE OF THE REGRESSION COEFFICIENT

Suppose the null-hypothesis to be tested specifies that the population value of the regression coefficient of y on x is β_{21}. In the case when the marginal variances are known, we use the statistic

$$t' = \left(\frac{\kappa_{20}}{\kappa_{02}}\right)^{\frac{1}{2}} \frac{\sqrt{(N-1)}}{\sqrt{(1-\rho^2)}} (b_{21}-\beta_{21}) \tag{27}$$

for the purpose, which follows Student's t-distribution with $N-1$ degrees of freedom in the normal-theory case.

In the case of unknown marginal variances, we use the statistic

$$t = \frac{\sqrt{(N-2)}\,(m_{11}-m_{20}\beta_{21})}{\sqrt{(m_{20}m_{02}-m_{11}^2)}}, \tag{28}$$

which is known to be distributed as Student's t with $N-2$ degrees of freedom in the case of a normal parent population.

It is straightforward to derive the distribution of t' from the frequency function of b_{21} expressed by (12). The derived expression is independent of κ_{20} and κ_{02}, but owing to its lengthy form, we do not propose to give it here.

We proceed to derive the frequency distribution of t defined by (28). In doing so, it may appear unsound to assume that the higher-order population cumulants are known while the second-order moments are not. But as has been remarked earlier, here our objective is to find roughly how much non-normality can be allowed in the parent population without seriously affecting the normal theory results, and the derived distribution of t will be quite useful for the purpose.

Let us start with the joint distribution of the second-order moments denoted by (8), and consider its typical term

$$f_0(m_{20}, m_{11}, m_{02})\, M_{20}^{r_1} M_{11}^{r_2} M_{02}^{r_3} (m_{20}m_{02}-m_{11}^2)^s, \tag{29}$$

where r_1, r_2, r_3 take the values 0, 1, 2, 3 such that $r = r_1+r_2+r_3 \leqslant 3$ and s takes the values 0, 1.

Putting

$$m_{02} = \frac{N-2}{m_{20}t^2}(m_{11}-m_{20}\beta_{21})^2 + \frac{m_{11}^2}{m_{20}} \tag{30}$$

in (29), we have to integrate our m_{11} and m_{20} in order to get the corresponding frequency function of t. This can be done by making the substitution

$$\theta = \frac{\kappa_{20}}{(1-\rho^2)^2\kappa_{02}m_{20}^2}(m_{11}-m_{20}\beta_{21})^2, \tag{31}$$

and evaluating the integrals. We find

$$g_{r_1r_2r_3s}(t) = \frac{N^{N-1}(N-2)^{\frac{1}{2}(N-2)+s}\kappa_{02}^{s}(1-\rho^2)^{\frac{1}{2}(N-1)+r+2s}}{4\pi\Gamma(N-2)\,\kappa_{20}^{N-1+r+s}t^{N-1+2s}T^{r}}\int_0^\infty \exp\left(-\frac{Nm_{20}}{2\kappa_{20}}\right)m_{20}^{N-2+r+2s}$$
$$\times\int_0^\infty \exp\left[-\frac{N(1-\rho^2)}{2\kappa_{20}T}m_{20}\theta\right]\theta^{\frac{1}{2}(N-3)+s+r_3}(\rho^2\theta-2\rho T\theta^{\frac{1}{2}}+T)^{r_1}(\rho\theta-T\theta^{\frac{1}{2}})^{r_2}\,d\theta\,dm_{20} \tag{32}$$
$$= \frac{2^{r+2s}(\kappa_{20}\kappa_{02})^{s}(1-\rho^2)^{r+s}}{N^{r+2s}T^{r}[1+t^2/(N-2)]^{s}}g_0(t)$$
$$\times\sum_{k=0}^{r}\frac{\Gamma[\frac{1}{2}(N-1)+s+\frac{1}{2}k]\,\Gamma[\frac{1}{2}(N-1)+s+r-\frac{1}{2}k]}{\{\Gamma[\frac{1}{2}(N-1)]\}^2}a_k(\rho,T)\frac{T^{\frac{1}{2}k}}{(1-\rho^2)^{\frac{1}{2}k}}, \tag{32 bis}$$

where
$$T = \frac{t^2}{t^2+N-2}, \tag{33}$$

$g_0(t)$ is the frequency function of Student's t with $N-2$ degrees of freedom, and $a_k(\rho,T)$ is given by

$$(\rho^2\theta-2\rho T\theta^{\frac{1}{2}}+T)^{r_1}(\rho\theta-T\theta^{\frac{1}{2}})^{r_2}\theta^{r_3} = \sum_{k=0}^{r}a_k(\rho,T)\,\theta^{\frac{1}{2}k}. \tag{34}$$

Using (32) *bis*, after some simplification, we can write the frequency function of t as obtained from (8) in the form

$$g(t) = g_0(t)\left[1+\frac{N-1}{2N(N+1)(1-\rho^2)^{\frac{1}{2}}}\left\{(N-1)\sqrt{(1-\rho^2)}\,\alpha_{11}[(N-1)T-1]+2\left(\frac{\Gamma(\frac{1}{2}N)}{\Gamma[\frac{1}{2}(N-1)]}\right)^2\alpha_{12}T^{\frac{1}{2}}\right\}\right.$$
$$-\frac{N-2}{N(N+1)(N+3)(1-\rho^2)^{\frac{1}{2}}}\left\{(N-1)\sqrt{(1-\rho^2)}\left[2\alpha_{21}\{(N-1)T-1\}+(N-1)\alpha_{22}\left(1-\frac{1}{t^2}\right)T\right]\right.$$
$$\left.\left.+\left(\frac{\Gamma(\frac{1}{2}N)}{\Gamma[\frac{1}{2}(N-1)]}\right)^2T^{\frac{1}{2}}\left[3\alpha_{23}+4N(1-\rho^2)\alpha_{24}\left(\frac{N}{3}T-1\right)+\frac{N^2}{3}(1-\rho^2)\alpha_{25}\left(1-\frac{3}{t^2}\right)T\right]\right\}\right], \tag{35}$$

where
$$\left.\begin{aligned}
\alpha_{11} &= \rho^2\lambda_{40}-2\rho\lambda_{31}+\lambda_{22},\\
\alpha_{12} &= \rho\lambda_{40}-(1+2\rho^2)\lambda_{31}+3\rho\lambda_{22}-\lambda_{13},\\
\alpha_{21} &= \rho^2\lambda_{30}^2-2\rho(1+\rho^2)\lambda_{30}\lambda_{21}+2\rho^2\lambda_{30}\lambda_{12}+(1+3\rho^2)\lambda_{21}^2-4\rho\lambda_{21}\lambda_{12}+\lambda_{12}^2,\\
\alpha_{22} &= \lambda_{30}\lambda_{12}-\lambda_{21}^2-\rho\lambda_{30}\lambda_{03}+\rho\lambda_{21}\lambda_{12}-\lambda_{12}^2+\lambda_{21}\lambda_{03},\\
\alpha_{23} &= \rho\lambda_{30}^2-(1+4\rho^2)\lambda_{30}\lambda_{21}+2\rho(1+\rho^2)\lambda_{30}\lambda_{12}+2\rho(2+\rho^2)\lambda_{21}^2\\
&\quad -\rho^2\lambda_{30}\lambda_{03}-(2+7\rho^2)\lambda_{21}\lambda_{12}+3\rho\lambda_{12}^2+2\rho\lambda_{21}\lambda_{03}-\lambda_{12}\lambda_{03},\\
\alpha_{24} &= \rho^3\lambda_{30}^2-3\rho^2\lambda_{30}\lambda_{21}+\rho\lambda_{30}\lambda_{12}+2\rho\lambda_{21}^2-\lambda_{21}\lambda_{12},\\
\alpha_{25} &= 2\rho\lambda_{30}\lambda_{12}-2\rho\lambda_{21}^2-\lambda_{30}\lambda_{03}+\lambda_{21}\lambda_{12}.
\end{aligned}\right\} \tag{36}$$

The distribution of t is again independent of κ_{20} and κ_{02}. The upper tail area of the distribution is given by

$$P\{t\geqslant t_0\} = \int_{t_0}^{\infty}g(t)\,dt$$
$$= \tfrac{1}{2}[1-I_{u_0}(\tfrac{1}{2},\tfrac{1}{2}N-1)]+\frac{(N-1)(1-u_0)^{\frac{1}{2}N-1}}{2N(N+1)(1-\rho^2)^{\frac{1}{2}}}\left[\frac{(N-1)\sqrt{(1-\rho^2)}\,u_0^{\frac{1}{2}}}{B(\frac{1}{2},\frac{1}{2}N-1)}\alpha_{11}+\frac{1}{B(\frac{1}{2},\frac{1}{2}(N-1))}\alpha_{12}\right]$$
$$-\frac{(N-2)(1-u_0)^{\frac{1}{2}N-1}}{N(N+1)(N+3)(1-\rho^2)^{\frac{1}{2}}}\left[\frac{(N-1)\sqrt{(1-\rho^2)}\,u_0^{\frac{1}{2}}}{B(\frac{1}{2},\frac{1}{2}N-1)}\left(\frac{N-1}{N-2}\alpha_{22}+2\alpha_{21}\right)\right.$$
$$\left.+\frac{1}{2B(\frac{1}{2},\frac{1}{2}(N-1))}\left\{3\alpha_{23}+\frac{N(1-\rho^2)}{3}\{4[(N-2)u_0-1]\alpha_{24}+[(N+1)u_0-1]\alpha_{25}\}\right\}\right], \tag{37}$$

where $u_0 = t_0^2/(t_0^2+N-2)$ and I_{u_0} denotes the incomplete beta function.

It is of interest to see here that as N tends to infinity, the corrective terms do not all disappear, and the formula (37) takes the form

$$P\{t \geqslant t_0\} = \frac{1}{\sqrt{(2\pi)}} \int_{t_0}^{\infty} e^{-\frac{1}{2}t^2}\, dt + \frac{\alpha_{11} t_0}{2(1-\rho^2)\sqrt{(2\pi)}}\, e^{-\frac{1}{2}t_0^2}. \tag{38}$$

This shows that for large N, only ρ and the coefficient α_{11} ($= \rho^2\lambda_{40} - 2\rho\lambda_{31} + \lambda_{22}$) will be predominant in contributing towards the effect of non-normality. Thus ordinarily for a given t_0, the contribution of positive α_{11} will be to increase the normal theory upper tail-area and that of negative α_{11} will be to decrease it.

Generally, a two-sided test with equal tail areas is used when the null hypothesis simply specifies the value of β_{21} and the alternatives can be of any form. To examine how the tail areas are affected by parental non-normality in such a case, we may use the formula

$$\begin{aligned} P\{|t| \geqslant t_0\} &= \int_{-\infty}^{-t_0} g(t)\, dt + \int_{t_0}^{\infty} g(t)\, dt \\ &= [1 - I_{u_0}(\tfrac{1}{2}, \tfrac{1}{2}N - 1)] + \frac{(N-1)\, u_0^{\frac{1}{2}}(1-u_0)^{\frac{1}{2}N-1}}{N(N+1)(1-\rho^2)\, B(\frac{1}{2}, \frac{1}{2}N-1)} \\ &\quad \times \left[(N-1)\alpha_{11} - \frac{2(N-2)}{(N+3)(1-\rho^2)}\left\{2\alpha_{21} + \frac{N-1}{N-2}\alpha_{22}\right\}\right]. \end{aligned} \tag{39}$$

5. FURTHER WORK

It is proposed in a continuation of this paper to study numerically the order of the effects which the various cumulants of the population will have on the mean and the variance of b_{21} and on the probability points of t. A comparison of these with known results will also be made. A considerable part of the computation involved is already finished and it is hoped to be able to publish the complete results in the next issue of *Biometrika*.

I should like to acknowledge my indebtedness to Dr A. K. Gayen for his kind advice and help in the course of preparation of this paper and also to Prof. E. S. Pearson from whom I have received helpful criticism on certain aspects of this work.

REFERENCES

COOK, M. B. (1951). *Biometrika*, **38**, 368.
GAYEN, A. K. (1951). *Biometrika*, **38**, 219.
HILL, I. D. (1954). *Biometrika*, **41**, 548.
HYRENIUS, H. (1952). *Biometrika*, **39**, 238.
PEARSON, E. S. (1929). *Biometrika*, **21**, 337.
QUENSEL, C. E. (1938). *Lunds. Univ. Årss.* N.F. Adv. 2, Bd. 34, Nr. 4; *Kungl. Fys. Sal. Handl.* N.F. Bd. 49, Nr. 4.
WICKSELL, S. D. (1917). *Kungl. Sr. Vet. Akad. Handl.* Bd. 58, no. 3, 1–48.
WICKSELL, S. D. (1933). *Biometrika*, **25**, 121.

Biometrika (1960), **47**, 1 *and* 2, *p.* 69
Printed in Great Britain

The separation of molecular compounds by countercurrent dialysis: a stochastic process

BY MARVIN A. KASTENBAUM

*Mathematics Panel and Biology Division, Oak Ridge National Laboratory**

1. BACKGROUND

Peptides isolated from rat-liver microsomes may be separated from contaminating amino acids, sugars and salts by countercurrent dialysis. A single such dialysis can achieve only fractional purification. Therefore, more complex systems of multiple dialysis, which achieve a higher degree of purification, have been devised.

The system with which this paper will be concerned is one originally suggested by Craig and associates (Craig & King, 1955, 1956; Craig, King & Stracher, 1957), in which a linear series of cells is used. Each dialysis cell is termed a stage, and each dialysis period is termed a cycle. After each cycle, the dialout† at each stage is concentrated and used as the dialin† in the succeeding stage, whereas the dialin at each stage is concentrated and returned to the dialysis sac at the previous stage. The dialin at the first stage is retained at the first stage; and the dialout at the final stage is taken out of the system. In the nomenclature of stochastic processes, the first stage represents a reflecting barrier and the last stage an absorbing barrier. The problem is to determine the probability with which a particle of the isolate will be at a specified stage in the system after a given number of cycles have been carried out.

2. FORMULATION

2·1. *The diffusion matrix*

Consider an m-stage system. Let p be the probability that, at any stage, a particle inside the dialysis sac will not penetrate to the outside volume, and $(1-p)$ the probability that it will penetrate to the outside. Also specify that, once outside, the particle cannot return inside, except by manual transfer. Then the matrix of diffusion probabilities for m stages is the square matrix of order $2m$ given by

$$D = \begin{bmatrix} D_1 & 0 & \dots & 0 \\ 0 & D_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & D_m \end{bmatrix}, \tag{2·1}$$

where

$$D_j = \begin{bmatrix} p & (1-p) \\ 0 & 1 \end{bmatrix} \quad (\text{for all } j = 1, 2, \dots, m)$$

is the matrix of diffusion probabilities for the jth stage. The odd rows and columns of D correspond to events taking place inside the dialysis sac for successive stages and the even rows and columns correspond to events taking place on the outside for successive stages.

* Operated by Union Carbide Corporation for the U.S. Atomic Energy Commission.

† These terms have been suggested by Anderson & Pennington (1959) to dispel the ambiguity in the term dialysate. Convenient abbreviations for the solutions inside and outside the dialysis sac are 'dialin' and 'dialout', respectively.

2·2. *The transfer matrix*

A cycle is defined as a constant interval of time commencing with the transfer of material into the dialysis sac (the beginning of the diffusion process) and concluding at the end of the diffusion process.

The transfer matrix, T, is a square matrix of order $2m$ whose elements are probabilities equal to zero or unity and which may be constructed according to the following scheme: At the end of each cycle

(i) a particle on the inside at stage 1 will remain on the inside of stage 1 with probability 1;

(ii) a particle on the inside at stage j will be transferred to the inside of stage $(j-1)$ with probability 1, for $j = 2, 3, \ldots, m$;

(iii) a particle on the outside of stage j will be transferred to the inside of stage $(j+1)$ with probability 1, for $j = 1, 2, \ldots, (m-1)$;

(iv) a particle on the outside of stage m will remain on the outside of stage m with probability 1.

The system may be completely described by recursion formulas as follows: Let P_{ij} be the probability that a particle is in the inside volume of the jth stage at the end of the ith cycle, and $(1-p)P_{ij}/p$ the probability that it is in the outside volume, where, for all cycles and stages $0 \leqslant p \leqslant 1$ is the probability that a particle inside will stay inside, $(1-p)$ is the probability that a particle inside will penetrate to the outside volume, and where the probability is zero that a particle will pass from the outside volume to the inside volume.

Then the probability that a particle is in the inside volume of stage $j = 1, 2, \ldots, m$ at the end of cycle $i = 1, 2, \ldots, n$ is given by the recursions

$$P_{i1} = p[P_{i-1,1} + P_{i-1,2}],$$

$$P_{ij} = (1-p)P_{i-1,j-1} + pP_{i-1,j+1} \quad [\text{for } j = 2, 3, \ldots, (m-1)],$$

and $$P_{im} = (1-p)P_{i-1,m-1},$$

where $$P_{01} = 1, \quad P_{0j} = 0 \quad (\text{for } j = 2, 3, \ldots, m),$$

and $$P_{ij} = 0 \quad (\text{for } i < j).$$

2·3. *The stochastic matrix*

If D and T are two square matrices as defined, then their product DT is a square matrix of order $2m$ with non-negative elements and unit row sums. The elements of this matrix are the conditional probabilities that a particle will make a transition from one of the $2m$ states to another. Such a matrix is a stochastic matrix.

Then if $i = 1, 2, \ldots, n$ is the number of cycles through which the system is to be carried, $(DT)^i$ is the matrix of conditional probabilities at the beginning of cycle $(i+1)$. By the same token, $(DT)^i D$ is the matrix of conditional probabilities at the end of cycle $(i+1)$, before transfer has been carried out.

If, as is the case in this system, all the material to be dialysed is placed in the dialysis sac at the first stage, the initial probabilities associated with the $2m$ states are represented by the $2m$-dimensional unit vector

$$e_1 = (1, 0, 0, \ldots, 0). \tag{2·2}$$

The probabilities at the beginning and the end of cycle $(i+1)$ are then respectively

$$e_1(DT)^i \quad \text{and} \quad e_1(DT)^i D. \tag{2·3}$$

The discussion in the succeeding sections will involve the algebraic reduction of the matrix $(DT)^i$ for a three-stage system and then for the general m-stage system.

3. The matrix $(DT)^i$

By the rules for its construction, the transfer matrix, T, will have $(m-1)$ null columns, namely, all those columns, except the $2m$th, involving transfers to the outside volume. Hence the matrix DT as well as its power $(DT)^i$ will have $(m-1)$ null columns whose positions are known. This property suggests that an initial simplification would be to reduce the matrix DT of order $2m$ to a matrix $\mathbf{DT}$ of order $(m+1)$, by eliminating the $(m-1)$ null columns and their respective rows. The resulting matrix is of full rank, and is given as

$$\mathbf{DT} = \begin{bmatrix} p & (1-p) & 0 & 0 & \dots & 0 & 0 & 0 \\ p & 0 & (1-p) & 0 & \dots & 0 & 0 & 0 \\ 0 & p & 0 & (1-p) & \dots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \dots & 0 & (1-p) & 0 \\ 0 & 0 & 0 & 0 & \dots & p & 0 & (1-p) \\ 0 & 0 & 0 & 0 & \dots & 0 & 0 & 1 \end{bmatrix}. \tag{3·1}$$

Let Λ be the Jordan normal form of $\mathbf{DT}$. Then we may write

$$(\mathbf{DT})^i = M^{-1}\Lambda^i M, \tag{3·2}$$

where M is a square matrix of order $(m+1)$ of characteristic vectors resulting from the solutions of the equations

$$\underline{x}[\mathbf{DT} - \lambda I] = 0 \tag{3·3}$$

and where $\underline{x}$ is the $(m+1)$-dimensional invariant row vector $(x_1, x_2, \dots, x_{m+1})$. In this instance, Λ may be shown to be a diagonal matrix of order $(m+1)$ whose elements are the roots of the determinantal equation

$$|\mathbf{DT} - \lambda I| = 0. \tag{3·4}$$

Thus there will always be one known root equal to unity, and m unknown roots whose values may be determined from the solution of the mth degree polynomial equation in λ

$$\psi_m(\lambda) = (-1)^m \begin{vmatrix} p-\lambda & 1-p & 0 & 0 & \dots & 0 & 0 & 0 \\ p & -\lambda & 1-p & 0 & \dots & 0 & 0 & 0 \\ 0 & p & -\lambda & 1-p & \dots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \dots & -\lambda & 1-p & 0 \\ 0 & 0 & 0 & 0 & \dots & p & -\lambda & 1-p \\ 0 & 0 & 0 & 0 & \dots & 0 & p & -\lambda \end{vmatrix} = 0. \tag{3·5}$$

4. The reduction of $(\mathbf{DT})^i$ for $m = 3$

The diffusion and transfer matrices for a three-stage system are each square matrices of order six and may be written respectively as

$$D = \begin{bmatrix} p & (1-p) & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & p & (1-p) & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & p & (1-p) \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \tag{4·1}$$

and

$$T = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}. \tag{4·2}$$

Their product DT is of order six and of rank four, containing two null column vectors in the second and fourth positions. By eliminating the two null columns and their respective rows, we form the matrix

$$\mathbf{DT} = \begin{bmatrix} p & 1-p & 0 & 0 \\ p & 0 & 1-p & 0 \\ 0 & p & 0 & 1-p \\ 0 & 0 & 0 & 1 \end{bmatrix}. \tag{4·3}$$

The characteristic equation (3·4) yields one root equal to unity and three unknown roots λ_1, λ_2 and λ_3 whose values may be determined from the polynomial equation

$$\psi_3(\lambda) = \lambda^3 - p\lambda^2 - 2p(1-p)\lambda + p^2(1-p) = 0. \tag{4·4}$$

It follows from (4·4) that

$$\left.\begin{aligned} \lambda_1+\lambda_2+\lambda_3 &= p = \sigma_1, \\ \lambda_1\lambda_2+\lambda_1\lambda_3+\lambda_2\lambda_3 &= -2p(1-p) = \sigma_2, \\ \text{and} \qquad \lambda_1\lambda_2\lambda_3 &= -p^2(1-p) = \sigma_3, \end{aligned}\right\} \tag{4·5}$$

whence it is easy to show that

$$(\lambda_1-1)(\lambda_2-1)(\lambda_3-1) = -(1-p)^3.$$

The matrix of characteristic vectors, M, is formed by solving for $\underline{x}$ in equation (3·3), when $m = 3$. Equation (3·3) in this case reduces to

$$\left.\begin{aligned} (1-p)x_3+(1-\lambda)x_4 &= 0, \\ (1-p)x_2-\lambda x_3 &= 0, \\ (1-p)x_1-\lambda x_2+px_3 &= 0 \\ \text{and} \qquad (p-\lambda)x_1+px_2 &= 0. \end{aligned}\right\} \tag{4·6}$$

When $\lambda = 1$, $x_1 = x_2 = x_3 = 0$, and we set $x_4 = 1$. For $\lambda = \lambda_j \neq 1$ $(j = 1, 2, 3)$, set

$$x_4 = (1-p)^2/(\lambda_j - 1). \tag{4·7}$$

Then

$$\left.\begin{aligned} x_3 &= \frac{(\lambda_j-1)x_4}{1-p} = (1-p), \\ x_2 &= \frac{\lambda_j x_3}{1-p} = \lambda \\ \text{and} \qquad x_1 &= \frac{\lambda_j x_2 - px_3}{1-p} = \frac{\lambda_j^2}{1-p} - p, \end{aligned}\right\} \tag{4·8}$$

and the matrix of characteristic vectors may then be written as

$$M = \begin{bmatrix} \frac{\lambda_1^2}{1-p} - p & \lambda_1 & (1-p) & \frac{(1-p)^2}{\lambda_1 - 1} \\ \frac{\lambda_2^2}{1-p} - p & \lambda_2 & (1-p) & \frac{(1-p)^2}{\lambda_2 - 1} \\ \frac{\lambda_3^2}{1-p} - p & \lambda_3 & (1-p) & \frac{(1-p)^2}{\lambda_3 - 1} \\ 0 & 0 & 0 & 1 \end{bmatrix}. \tag{4·9}$$

The determinant of M

$$|M| = \begin{vmatrix} \frac{\lambda_1^2}{1-p} - p & \lambda_1 & (1-p) \\ \frac{\lambda_2^2}{1-p} - p & \lambda_2 & (1-p) \\ \frac{\lambda_3^2}{1-p} - p & \lambda_3 & (1-p) \end{vmatrix} = \begin{vmatrix} \lambda_1^2 & \lambda_1 & 1 \\ \lambda_2^2 & \lambda_2 & 1 \\ \lambda_3^2 & \lambda_3 & 1 \end{vmatrix} = \Delta. \tag{4·10}$$

Then

$$\epsilon_1 M^{-1} = \frac{1}{\Delta}\left[(1-p)\begin{vmatrix} \lambda_2 & 1 \\ \lambda_3 & 1 \end{vmatrix}, \quad -(1-p)\begin{vmatrix} \lambda_1 & 1 \\ \lambda_3 & 1 \end{vmatrix}, \quad (1-p)\begin{vmatrix} \lambda_1 & 1 \\ \lambda_2 & 1 \end{vmatrix}, \quad -(1-p)^3\begin{vmatrix} \lambda_1 & 1 & \frac{1}{\lambda_1 - 1} \\ \lambda_2 & 1 & \frac{1}{\lambda_2 - 1} \\ \lambda_3 & 1 & \frac{1}{\lambda_3 - 1} \end{vmatrix}\right]$$

$$= \frac{1}{\Delta}[(1-p)V_{11},\ (1-p)V_{21},\ (1-p)V_{31},\ \Delta], \tag{4·11}$$

where ϵ_1 is the four-dimensional vector $(1, 0, 0, 0)$, and V_{j1} is the co-factor of the element in the jth row and first column of Δ. Set

$$\Lambda = \begin{bmatrix} \lambda_1 & 0 & 0 & 0 \\ 0 & \lambda_2 & 0 & 0 \\ 0 & 0 & \lambda_3 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

Then $$\epsilon_1 M^{-1}\Lambda^i = \left[\frac{(1-p)\lambda_1^i V_{11}}{\Delta}, \frac{(1-p)\lambda_2^i V_{21}}{\Delta}, \frac{(1-p)\lambda_3^i V_{31}}{\Delta}, 1\right]$$

and

$$\epsilon_1 M^{-1}\Lambda^i M = \epsilon_1(\mathbf{DT})^i = \left[\frac{1}{\Delta}\{R_{i+2} - p(1-p)R_i\}, \frac{(1-p)}{\Delta}R_{i+1}, \frac{(1-p)^2}{\Delta}R_i, 1 - \frac{(1-p)^3}{\Delta}U_i\right], \tag{4·12}$$

where $$R_i = \sum_{j=1}^{3} \lambda_j^i V_{j1} \quad \text{and} \quad U_i = -\sum_{j=1}^{3} \frac{\lambda_j^i}{\lambda_j - 1} V_{j1}.$$

It follows that

$$R_0 = \begin{vmatrix} 1 & \lambda_1 & 1 \\ 1 & \lambda_2 & 1 \\ 1 & \lambda_3 & 1 \end{vmatrix} = 0, \qquad R_1 = \begin{vmatrix} \lambda_1 & \lambda_1 & 1 \\ \lambda_2 & \lambda_2 & 1 \\ \lambda_3 & \lambda_3 & 1 \end{vmatrix} = 0, \qquad R_2 = \begin{vmatrix} \lambda_1^2 & \lambda_1 & 1 \\ \lambda_2^2 & \lambda_2 & 1 \\ \lambda_3^2 & \lambda_3 & 1 \end{vmatrix} = \Delta,$$

and $$R_i = pR_{i-1} + 2p(1-p)R_{i-2} - p^2(1-p)R_{i-3} \quad (\text{for } i = 3, 4, \ldots, n). \tag{4·13}$$

Note that the coefficients in the recursion (4·13) are identical with those of the polynomial (4·4). Also

$$U_0 = -\begin{vmatrix} \frac{1}{\lambda_1 - 1} & \lambda_1 & 1 \\ \frac{1}{\lambda_2 - 1} & \lambda_2 & 1 \\ \frac{1}{\lambda_3 - 1} & \lambda_3 & 1 \end{vmatrix} = \frac{\Delta}{(1-p)^3}$$

and, in general, $$U_i = U_{i-1} - R_{i-1} \quad (\text{for } i = 1, 2, \ldots, n). \tag{4·14}$$

Thus the vector (4·12) may be evaluated for any i by applying the recursions expressed in (4·13) and (4·14).

The six-dimensional vector of probabilities at the beginning of cycle $(i+1)$ may be written directly from (4·12) by replacing two zero elements in the same positions from which the null column vectors were previously dropped. Thus

$$e_1(DT)^i = \left[\frac{1}{\Delta}\{R_{i+2} - p(1-p)R_i\},\ 0,\ \frac{(1-p)}{\Delta}R_{i+1},\ 0,\ \frac{(1-p)^2}{\Delta}R_i,\ 1 - \frac{(1-p)^3}{\Delta}U_i\right] \tag{4·15}$$

and

$$e_1(DT)^i D = \left[\frac{p}{\Delta}\{R_{i+2} - p(1-p)R_i\},\ \frac{(1-p)}{\Delta}\{R_{i+2} - p(1-p)R_i\},\ \frac{p(1-p)}{\Delta}R_{i+1},\right.$$
$$\left.\frac{(1-p)^2}{\Delta}R_{i+1},\ \frac{p(1-p)^2}{\Delta}R_i,\ 1 - \frac{(1-p)^3}{\Delta}U_{i+1}\right]. \tag{4·16}$$

5. THE mTH ORDER POLYNOMIAL IN λ

If $\psi_m(\lambda)$ be the mth order polynomial in λ, then the recursion

$$\psi_m(\lambda) = \lambda\psi_{m-1}(\lambda) - p(1-p)\psi_{m-2}(\lambda) \tag{5·1}$$

holds, with the initial conditions that

$$\psi_0(\lambda) = 1 \quad \text{and} \quad \psi_1(\lambda) = \lambda - p. \tag{5·2}$$

A particular solution of (5·1) is $$y^m = \psi_m(\lambda) \tag{5·3}$$

if y is either y_1 or y_2, where y_1 and y_2 are the roots of the quadratic

$$y^2 - \lambda y + p(1-p) = 0. \tag{5·4}$$

Moreover, any solution of the recursion is a linear combination of any pair of independent solutions. But $\psi_m(\lambda)$ is the particular solution whose values are 1 and $\lambda - p$ at 0 and 1, respectively. Hence, we may write

$$\psi_m(\lambda) = B_1 y_1^m + B_2 y_2^m, \tag{5·5}$$

whence it follows that

$$B_1 + B_2 = 1 \quad \text{and} \quad B_1 y_1 + B_2 y_2 = \lambda - p. \tag{5·6}$$

The simultaneous solution of equations (5·6) is

$$B_1 = \frac{(\lambda - p) - y_2}{y_1 - y_2} \quad \text{and} \quad B_2 = \frac{y_1 - (\lambda - p)}{y_1 - y_2}, \tag{5·7}$$

which, when substituted into (5·5), yields

$$\psi_m(\lambda) = \frac{1}{y_1 - y_2}\{(y_1^{m+1} - y_2^{m+1}) - p(y_1^m - y_2^m)\}. \tag{5·8}$$

Now let $y_1 = r e^{i\theta}$ and $y_2 = r e^{-i\theta}$,

where $r = \sqrt{[p(1-p)]}$ and $\cos\theta = -\dfrac{\lambda}{2r}$, $\sin\theta > 0$.

Then $$\psi_m(\lambda) = \frac{r^m \sin(m+1)\theta - p r^{m-1}\sin m\theta}{\sin\theta}. \tag{5·9}$$

If $t = \cos\theta$, then $$C_k(t) = \frac{\sin(k+1)\theta}{\sin\theta} \tag{5·10}$$

is the kth-order Chebyshev polynomial of the second kind. Thus one method of evaluating the roots of the mth-order polynomial in λ would be to solve for $t = -\frac{1}{2}\lambda\sqrt{[p(1-p)]}$ in the equation

$$C_m(t) - \sqrt{\left(\frac{p}{1-p}\right)} C_{m-1}(t) = 0, \tag{5·11}$$

or $$\frac{C_{m-1}(t)}{C_m(t)} = \sqrt{\left(\frac{1-p}{p}\right)}. \tag{5·12}$$

The polynomials C_k can be obtained directly by employing the methods outlined by Householder (1958).

Though explicit values of the unknown roots are not available in general from (5·11) or from any other solution of the mth-order polynomial equation, the elementary symmetric functions of these roots, derived from the coefficients of the polynomial, may be used in evaluating (2·3) explicitly. The coefficients of the mth degree polynomial may be obtained formally from (5·1) by use of generating functions in the following way:* Write

$$\psi_m(\lambda) = b_0^m + b_1^m\lambda + \ldots + b_m^m\lambda^m \tag{5·13}$$

and $$\psi(z,\lambda) = \sum_{m=0}^{\infty} \psi_m(\lambda) z^m = \sum_{k=0}^{\infty} \lambda^k \sum_{j=k}^{\infty} b_k^j z^j. \tag{5·14}$$

From (5·1) and (5·2) we obtain

$$\psi(z,\lambda) = \frac{1-pz}{1-\lambda z + p(1-p)z^2} = \frac{1-pz}{1+p(1-p)z^2}\left[1 - \frac{\lambda z}{1+p(1-p)z^2}\right]^{-1}. \tag{5·15}$$

Expanding (5·15) in powers of λ, and letting $q = p(1-p)$, we obtain

$$\psi(z,\lambda) = \sum_{k=0}^{\infty} \frac{(1-pz)z^k\lambda^k}{[1+qz^2]^{k+1}}. \tag{5·16}$$

Thus from (5·14) we may write

$$\sum_{j=k}^{\infty} b_k^j z^j = (1-pz)z^k\left[1 - \binom{k+1}{1} qz^2 + \binom{k+2}{2} q^2 z^4 - \ldots\right], \tag{5·17}$$

whence $$b_k^{k+2s} = (-1)^s \binom{k+s}{s} [p(1-p)]^s \tag{5·18}$$

and $$b_k^{k+2s+1} = (-1)^{s+1} \binom{k+s}{s} p^{s+1}(1-p)^s, \quad \text{for} \quad k = 0, 1, 2, \ldots. \tag{5·19}$$

* The author wishes to express his gratitude to the referee who, in addition to making other helpful suggestions, provided the method for directly evaluating the coefficients of $\psi_m(\lambda)$.

Before passing to the ultimate reduction of $(\mathbf{DT})^i$ it might be worth noting that the roots of the mth-order polynomial are all real and distinct. That they are real may be demonstrated by a simple contragredient transformation on **DT**, which reduces it to a triple diagonal symmetric matrix; that they are both real and distinct may be shown by a direct application of the argument presented by Householder (1953, pp. 181–2).

6. THE GENERAL REDUCTION OF $(\mathbf{DT})^i$ FOR ANY m

Consider the square matrix **DT** of order $(m+1)$ as given in (3·1), and let $\underline{x}$ be the characteristic vector as defined in (3·3). If (3·4) is the determinantal equation, then $\lambda = 1$ is always one root occurring with multiplicity 1. Associated with this root is the $(m+1)$-dimensional characteristic vector

$$\epsilon_{m+1} = (0, 0, \ldots, 0, 1). \tag{6·1}$$

In general, the elements of $\underline{x}$ when $\lambda = \lambda_j \neq 1$ (for $j = 1, 2, \ldots, m$) may be found from the simultaneous equations

$$\left.\begin{aligned} (1-p)\,x_m + (1-\lambda_j)\,x_{m+1} &= 0, \\ (1-p)\,x_{m-1} - \lambda_j x_m &= 0, \\ (1-p)\,x_k - \lambda_j x_{k+1} + p x_{k+2} &= 0 \quad (\text{for } k = 1, 2, \ldots, m-2), \\ \text{and} \qquad (p-\lambda_j)\,x_1 + p x_2 &= 0, \end{aligned}\right\} \tag{6·2}$$

by first setting

$$x_{m+1} = \frac{(1-p)^{m-1}}{\lambda_j - 1}. \tag{6·3}$$

Then

$$\left.\begin{aligned} x_m &= (1-p)^{m-2}, \\ x_{m-1} &= \lambda_j (1-p)^{m-3}, \\ \text{and} \qquad x_k &= \frac{\lambda_j x_{k+1} - p x_{k+2}}{(1-p)} \quad (\text{for } k = 1, 2, \ldots, m-2). \end{aligned}\right\} \tag{6·4}$$

Let M be the square matrix of order $(m+1)$ of characteristic vectors so generated. It may be demonstrated that the determinant of M

$$|M| = (1-p)^{\frac{1}{2}[m(m-3)]} \begin{vmatrix} \lambda_1^{m-1} & \lambda_1^{m-2} & \ldots & \lambda_1 & 1 \\ \lambda_2^{m-1} & \lambda_2^{m-2} & \ldots & \lambda_2 & 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \lambda_m^{m-1} & \lambda_m^{m-2} & \ldots & \lambda_m & 1 \end{vmatrix} \tag{6·5}$$

$$= (1-p)^{\frac{1}{2}[m(m-3)]} \prod_{j<k=1}^{m} (\lambda_j - \lambda_k) = (1-p)^{\frac{1}{2}[m(m-3)]}\,\Delta.$$

Then

$$\epsilon_1 M^{-1} = \frac{(1-p)}{\Delta}\left[V_{11}, V_{21}, \ldots, V_{m1}, \frac{\Delta}{(1-p)}\right], \tag{6·6}$$

where ϵ_1 is the $(m+1)$-dimensional vector $(1, 0, 0, \ldots, 0)$, and V_{j1} is the co-factor of the element in the jth row and first column in Δ, and

$$\epsilon_1 M^{-1} \Lambda^i = \left[\frac{(1-p)\,\lambda_1^i V_{11}}{\Delta}, \frac{(1-p)\,\lambda_2^i V_{21}}{\Delta}, \ldots, \frac{(1-p)\,\lambda_m^i V_{m1}}{\Delta}, 1\right] \tag{6·7}$$

where Λ is the Jordan normal form of **DT** as described in § 3.

Now define

$$R_i = \sum_{j=1}^{m} \lambda_j^i V_{j1} \quad (\text{for } i = 1, 2, \ldots, n). \tag{6·8}$$

Then, as in § 4 it is easy to see that

$$R_i = \begin{cases} 0 & [\text{for } i = 0, 1, 2, \ldots, (m-2)], \\ \Delta & [\text{for } i = (m-1)], \\ \sum_{s=1}^{m} (-1)^{s-1} \sigma_s R_{i-s} & [\text{for } i = m, (m+1), \ldots, n], \end{cases} \tag{6·9}$$

where σ_s is the sth elementary symmetric function of the roots of the polynomial $\psi_m(\lambda) = 0$. That is to say, the coefficients in the recursion (6·9) are identical with those of the polynomial equation $\psi_m(\lambda) = 0$.

Let

$$e_1 M^{-1} \Lambda^i M = (v_1, v_2, \ldots, v_{m+1}). \tag{6·10}$$

Then, if in equations (6·4), λ^i is replaced by R_{i+t},

$$v_j = \frac{(1-p)}{\Delta} x_j \quad (\text{for } j = 1, 2, \ldots, m), \tag{6·11}$$

and

$$v_{m+1} = 1 - \frac{(1-p)^m}{\Delta} U_i, \tag{6·12}$$

where

$$U_0 = \frac{\Delta}{(1-p)^m} \quad \text{and} \quad U_i = U_{i-1} - R_{i-1} \quad (\text{for } i = 1, 2, \ldots, n). \tag{6·13}$$

Finally

$$e_1 (DT)^i = (v_1, 0, v_2, 0, \ldots, 0, v_m, v_{m+1}) \tag{6·14}$$

and

$$e_1 (DT)^i D = [pv_1, (1-p)v_1, pv_2, (1-p)v_2, \ldots, pv_m, (1-p)v_m + v_{m+1}]. \tag{6·15}$$

References

Anderson, N. G. & Pennington, L. F. (1959). Studies on isolated cell components. XIV. Separation of peptides by countercurrent dialysis. To be submitted to *Exp. Cell Res.*

Craig, L. C. & King, T. P. (1955). Some dialysis experiments with polypeptides. *J. Amer. Chem. Soc.* **77**, 6620–4.

Craig, L. C. & King, T. P. (1956). Fractional dialysis with cellophane membranes. *J. Amer. Chem. Soc.* **78**, 4171–2.

Craig, L. C., King, T. P. & Stracher, A. (1957). Dialysis studies. II. Some experiments dealing with the problem of selectivity. *J. Amer. Chem. Soc.* **79**, 3729–37.

Householder, A. S. (1953). *Principles of Numerical Analysis.* New York, Toronto, and London: McGraw-Hill Book Co., Inc.

Householder, A. S. (1958). The approximate solution of matrix problems. *J. Ass. Comput. Machinery*, **5**, 205–43.

Biometrika (1960), **47**, 1 *and* 2, *p.* 79
Printed in Great Britain

A note on the error after a number of terms of the David–Johnson series for the expected values of normal order statistics

BY J. G. SAW
University College London

1. Introduction

David & Johnson (1954) proposed that a finite number of terms of a certain infinite series could be used as an approximation to the integral involved in evaluating the expected value of the rth smallest of n-ordered normal variates.

Plackett (1958) obtained a different infinite series and was able to give bounds for the error which resulted in using only a finite number of terms of this series.

It is the purpose of this note to derive bounds for the error which results in using only a finite number of terms of the David & Johnson series. With this accomplished, the last section is devoted to applying both Plackett's method and David & Johnson's method for a particular example so that the two techniques may be compared. Briefly, it is found that the convergence of the Plackett series is a little faster than the David & Johnson series and that the bounds which Plackett gives are slightly sharper than those which we shall derive for the David & Johnson series. However, these small differences seem to be negligible in the light of the great computational advantages of the David & Johnson technique over that of Plackett.

2. Notation

Let $x_{1:n} < x_{2:n} < \ldots < x_{n:n}$ be an ordered sample of n observations from the standard normal population.

Write

$$Z(x) = \frac{1}{\sqrt{(2\pi)}} \exp -\tfrac{1}{2}x^2, \quad P(x) = \int_{-\infty}^{x} Z(u)\,du, \quad p_r = 1 - q_r = r/(n+1).$$

We use $x(P)$ for the inverse function defined by

$$\int_{-\infty}^{x(P)} Z(u)\,du = P$$

and write

$$\beta_P(a,b) = \frac{1}{B(a,b)} P^{a-1}(1-P)^{b-1},$$

$$M_t(a,b) = \int_0^1 \left(P - \frac{a}{a+b}\right)^t \beta_P(a,b)\,dP.$$

$M_t(a,b)$ is thus the tth central moment of a Beta or Pearson Type I distribution and is hypergeometric in form.

We will use

$$x_t(P) \text{ to denote } \left(\frac{d}{dP}\right)^t x(P),$$

$$x_t(p_r) \text{ to denote } \left(\frac{d}{dP}\right)^t x(P)\Big|_{P=p_r}.$$

3. The David and Johnson expansion

The probability function of $x_{r:n}$ is given by

$$p(x_{r:n}) = \frac{1}{B(r, n-r+1)} \{P(x_{r:n})\}^{r-1} \{1 - P(x_{r:n})\}^{n-r} Z(x_{r:n}), \tag{3·1}$$

so that

$$\mathscr{E}(x_{r:n}) = \int_{-\infty}^{+\infty} p(x_{r:n})\, x_{r:n}\, dx_{r:n} \tag{3·2}$$

$$= \int_0^1 \beta_P(r, n-r+1)\, x(P)\, dP. \tag{3·3}$$

We have, following David & Johnson,

$$\mathscr{E}(x_{r:n}) = \sum_{t=0}^{2m} \frac{1}{t!} x_t(p_r)\, M_t(r, n-r+1) + R_{2m}. \tag{3·4}$$

R_{2m}, the remainder after $2m$ terms, is the difference between the true value of $\mathscr{E}(x_{r:n})$ and the sum of the series to $2m$ terms. In practice, the remainder is usually found to be small for quite small m so that the series may be used as an approximation to the true value of $\mathscr{E}(x_{r:n})$. Several forms for R_{2m} exist. We consider a particular one which may be obtained as follows.

Consider the Taylor series expansion for the function $x(P)$ about the point $P = p_r$; then

$$x(P) = \sum_{t=0}^{2m} \frac{1}{t!} x_t(p_r)\, (P - p_r)^t + R'_{2m}. \tag{3·5}$$

Using the Darboux form for this remainder (see, for example, C. A. Stewart's *Advanced Calculus*, p. 440), we write

$$R'_{2m} = \frac{1}{(2m)!} \int_0^1 (1-\gamma)^{2m}\, x_{2m+1}(p_r + \gamma(P - p_r))\, d\gamma. \tag{3·6}$$

Multiply both sides of equation (3·5) by $\beta_P(r, n-r+1)$ and integrate over P between zero and one. Using the definition of $M_t(r, n-r+1)$ and equation (3·3), we have exactly equation (3·4) with

$$R_{2m} = \int_0^1 \beta_P(r, n-r+1)\, R'_{2m}\, dP, \tag{3·7}$$

that is

$$R_{2m} = \frac{1}{(2m)!} \int_0^1 \int_0^1 (1-\gamma)^{2m}\, x_{2m+1}(p_r + \gamma(P - p_r))\, \beta_P(r, n-r+1)\, dP\, d\gamma. \tag{3·8}$$

We shall try to give upper and lower bounds for R_{2m}, using the form of equation (3·8).

4. The derivatives $x_j(P)$

We now discuss the behaviour of $x_j(P)$ and of

$$\eta_{2m+1}(P) = P^{2m+1} x_{2m+1}(P).$$

It is easily verified that for $j \geqslant 1$ we have

$$x_j(P) = Z^{-j}(x)\, H_j(x),$$

where $H_j(x)$ is a polynomial in x of degree $j-1$ and which is even or odd according as j is odd or even. Thus we have

$$H_1(x) = 1, \quad H_2(x) = x, \quad H_3(x) = 1 + 2x^2, \quad H_4(x) = 7x + 6x^3, \quad \text{etc.}$$

These polynomials obey the recurrence relation

$$H_{j+1}(x) = jxH_j(x) + H'_j(x) \tag{4·1}$$

and the author (1958) has given the numerical coefficients in $H_j(x)$ for $j = 1\,(1)\,10$. Further, it is easily shown that

$$x_j(P) = (-1)^{j-1} x_j(1-P). \tag{4·2}$$

Turning now to $\eta_j(P)$, we have immediately that $\eta_j(0) = 0$ and $\eta_j(P) \to \infty$ as $P \to 1$. For the later development we require to show that $\eta_{2m+1}(P)$ is an increasing function of P. This is obviously true in the range $\frac{1}{2} < P < 1$ since $\eta_{2m+1}(P)$ is the product of two increasing functions.

Suppose then that $0 < P < \frac{1}{2}$ (i.e. $-\infty < x < 0$) and we require to show that

$$\begin{aligned} 0 < \frac{d}{dP}\eta_{2m+1}(P) &= (2m+1)\,P^{2m}\,x_{2m+1}(P) + P^{2m+1}\,x_{2m+2}(P) \\ &= \frac{P^{2m}}{Z^{2m+2}(x)} H_{2m+2}(x)\left\{(2m+1)\,Z(x)\frac{H_{2m+1}(x)}{H_{2m+2}(x)} + P\right\} \\ &= F(P)\,G(P) \text{ say,} \end{aligned} \tag{4·3}$$

where $G(P)$ is the expression in brackets.

Now for $0 < P < \frac{1}{2}$, $F(P)$ is negative. Therefore since $G(0) = 0$, it is sufficient (but not necessary) that dG/dP or dG/dx be negative in this same range. Now

$$\frac{dG(P)}{dx} = \frac{-Z(x)\,A_{2m+1}(x)}{H^2_{2m+2}(x)}, \tag{4·4}$$

where

$$A_{2m+1}(x) = (2m+1)\,H_{2m+1}(x)\,H'_{2m+2}(x) - (2m+2)\,H'_{2m+1}(x)\,H_{2m+2}(x) \tag{4·5}$$

or, using equation (3·1),

$$A_{2m+1}(x) = (2m+1)\,H_{2m+1}(x)\,H_{2m+3}(x) - (2m+2)\,H^2_{2m+2}(x). \tag{4·6}$$

From equation (4·4) it is sufficient to show that $A_{2m+1}(x) > 0$ in the range $-\infty < x < 0$. I have been unable to show this in general; however, since the polynomials $H_j(x)$ are known explicitly it is a fairly simple matter to verify that $A_{2m+1}(x) > 0$ for given m. In fact, for the first few values of m, we find that $A_{2m+1}(x)$ is a polynomial of degree $2m$ in x^2, in which only the coefficient of x^2 is negative. Thus

$$\begin{aligned} A_1(x) &= 1, \\ A_3(x) &= 21 - 16x^2 + 12x^4, \\ A_5(x) &= 4445 - 6664x^2 + 8076x^4 + 1344x^6 + 2880x^8, \\ A_7(x) &= 3{,}884{,}041 - 8{,}6\text{[illegible]}{,}072x^2 + 14{,}468{,}040x^4 + 6{,}808{,}896x^6 + 24{,}184{,}656x^8 \\ &\quad + 11{,}093{,}760x^{10} + 3{,}628{,}800x^{12}, \\ A_9(x) &= 9{,}580{,}522{,}329 - 28{,}374{,}712{,}624x^2 + 60{,}246{,}981{,}384x^4 + 54{,}671{,}037{,}120x^6 \\ &\quad + 292{,}158{,}113{,}616x^8 + 337{,}984{,}717{,}824x^{10} + 259{,}516{,}161{,}792x^{12} \\ &\quad + 88{,}044{,}104{,}640x^{14} + 14{,}631{,}321{,}600x^{16}. \end{aligned}$$

In each case we note that $A_{2m+1}(x)$ is positive definite since the first three terms (i.e. in x^0, x^2 and x^4) form a positive definite quadratic in x^2 and hence the conjecture that $\eta_{2m+1}(P)$ is increasing as P increases from zero to one is proved for $m = 0, 1, 2, 3$ and 4.* Thus all cases likely to be required in practice are covered.

* I owe the suggestion of this proof to D. E. Barton.

Incidentally, since $\eta_{2m}(0) = \eta_{2m}(\frac{1}{2}) = 0$, it cannot be true that $\eta_{2m}(P)$ is an increasing function of P throughout the range $0 < P < 1$.

It is of interest to note that

$$A_j(x)\,Z(x) = jx_j(P)\,x_{j+2}(P) - (j+1)\,x_{j+1}^2(P)$$

and so using the complex integral form for these derivatives we have to show that

$$\int_C \frac{d\xi}{(P(\xi)-P(x))^j} \int_C \frac{d\xi}{(P(\xi)-P(x))^{j+2}} < \left\{\int_C \frac{d\xi}{(P(\xi)-P(x))^{j+1}}\right\}^2 \tag{4·7}$$

is true when j is odd for any x, but is not true for general x when j is even.

In equation (4·7), C is to be a closed contour including only the zero of $P(\xi)-P(x)$ which occurs at the point on the real axis at $\xi = x$.

5. Bounds for R_{2m}

It is easily shown that $\mathscr{E}(x_{r:n}) = -\mathscr{E}(x_{n-r+1:n})$. As a consequence we will consider only the cases for which $p_r > \frac{1}{2}$ and note that the results extend in an obvious manner to the cases for which $p_r < \frac{1}{2}$.

In equation (3·8) make the substitution

$$u = p_r + \gamma(P - p_r)$$

and we have

$$R_{2m} = \frac{1}{(2m)!}\left\{-\int_0^{p_r} du \int_0^u dP + \int_{p_r}^1 du \int_u^1 dP\right\} (P-u)^{2m}\, x_{2m+1}(u)\, \beta_P(r, n-r+1). \tag{5·1}$$

In the second of the two integrals, replacing

$$u \quad \text{by} \quad 1-u, \qquad P \quad \text{by} \quad 1-P$$

and using the symmetrical properties of $x_{2m+1}(u)$ expressed in equation (4·2), it is seen that

$$R_{2m} = \frac{1}{(2m)!}\iint_{0<P<u<q_r} (u-P)^{2m}\, x_{2m+1}(u)\{\beta_P(n-r+1, r) - \beta_P(r, n-r+1)\}\, du\, dP$$
$$- \frac{1}{(2m)!}\int_{q_r}^{p_r} du \int_0^u dP (u-P)^{2m}\, x_{2m+1}(u)\, \beta_P(r, n-r+1). \tag{5·2}$$

We write J_1 and J_2 for the first and second of the integrals occurring in equation (5·2) and observe that both integrals are positive, since the integrands are everywhere positive (when $q_r < \frac{1}{2}$). Now

$$J_1 = \frac{1}{(2m)!}\iint_{0<P<u<q_r} \left(1-\frac{P}{u}\right)^{2m} \frac{1}{u}\, \eta_{2m+1}(u)\{\beta_P(n-r+1, r) - \beta_P(r, n-r+1)\}\, du\, dP \tag{5·3}$$

and so, using the properties of $\eta_{2m+1}(u)$ described in § 4, we have

$$J_1 < \frac{\eta_{2m+1}(q_r)}{(2m)!}\iint_{0<P<u<q_r} \left(1-\frac{P}{u}\right)^{2m} \frac{1}{u}\{\beta_P(n-r+1, r) - \beta_P(r, n-r+1)\}\, du\, dP. \tag{5·4}$$

But

$$\int_P^{q_r} \left(1-\frac{P}{u}\right)^{2m} \frac{du}{u} < \int_P^{q_r} \left(1-\frac{P}{u}\right)^{2m} \frac{q_r}{u^2} du = \frac{(q_r - P)^{2m+1}}{(2m+1)\, q_r^{2m} P}.$$

Finally, therefore

$$J_1 < \frac{q_r x_{2m+1}(p_r)}{(2m+1)!}\int_0^{q_r}(q_r-P)^{2m+1}\frac{1}{P}\{\beta_P(n-r+1,r)-\beta_P(r,n-r+1)\}\,dP. \tag{5·5}$$

Returning to the definition of J_1 in equation (5·2), since $x_{2m+1}(u)$ is a decreasing function of u in $(0, q_r < \frac{1}{2})$, we have

$$J_1 > \frac{x_{2m+1}(p_r)}{(2m+1)!}\iint_{0<P<u<q_r}(u-P)^{2m+1}\{\beta_P(n-r+1,r)-\beta_P(r,n-r+1)\}\,dP\,du$$

so that, performing the integration over u,

$$J_1 > \frac{x_{2m+1}(p_r)}{(2m+1)!}\int_0^{q_r}(q_r-P)^{2m+1}\{\beta_P(n-r+1,r)-\beta_P(r,n-r+1)\}\,dP. \tag{5·6}$$

For J_2, since $x_{2m+1}(u)$ is an increasing function of x^2, we have

$$J_2 < \frac{x_{2m+1}(p_r)}{(2m)!}\int_{q_r}^{p_r}du\int_0^u dP(u-P)^{2m}\beta_P(r,n-r+1), \tag{5·7}$$

$$J_2 > \frac{x_{2m+1}(\frac{1}{2})}{(2m)!}\int_{q_r}^{p_r}du\int_0^u dP(u-P)^{2m}\beta_P(r,n-r+1). \tag{5·8}$$

Dividing the region of integration into two parts

$$\int_{q_r}^{p_r}du\int_0^u dP \equiv \int_{q_r}^{p_r}du\int_0^{q_r}dP+\int_{q_r}^{p_r}dP\int_p^{p_r}du$$

and integrating out over u we have, after some slight rearrangement of the integrals,

$$J_2 < \frac{-x_{2m+1}(p_r)}{(2m+1)!}\left\{M_{2m+1}(r,n-r+1)+\int_0^{q_r}(P-q_r)^{2m+1}\{\beta_P(n-r+1,r)-\beta_P(r,n-r+1)\}\,dP\right\}, \tag{5·9}$$

$$J_2 > \frac{-x_{2m+1}(\frac{1}{2})}{(2m+1)!}\left\{M_{2m+1}(r,n-r+1)+\int_0^{q_r}(P-q_r)^{2m+1}\{\beta_P(n-r+1,r)-\beta_P(r,n-r+1)\}\,dP\right\}. \tag{5·10}$$

Combining equations (5·6) and (5·9) we have after cancellation

$$R_{2m} > \frac{x_{2m+1}(p_r)}{(2m+1)!}M_{2m+1}(r,n-r+1). \tag{5·11}$$

Similarly, combining the results expressed in equations (5·5) and (5·10) we arrive at an upper bound for R_{2m} which may be conveniently written in the form

$$R_{2m} < \frac{q_r x_{2m+1}(p_r)}{(2m+1)!}\{V_{2m+1}(r,n-r+1:q_r)-V_{2m+1}(n-r+1,r:q_r)\}$$
$$-\frac{x_{2m+1}(\frac{1}{2})}{(2m+1)!}\{W_{2m+1}(r,n-r+1:q_r)-W_{2m+1}(n-r+1,r:q_r)-M_{2m+1}(r,n-r+1)\}, \tag{5·12}$$

where

$$W_s(a,b:\nu) = \int_0^{\nu}(P-\nu)^s\beta_P(a,b)\,dP, \tag{5·13}$$

$$V_s(a,b:\nu) = \int_0^{\nu}\frac{1}{P}(P-\nu)^s\beta_P(a,b)\,dP. \tag{5·14}$$

6. COMPUTATION OF $M_s(a, b)$, $W_s(a, b; \nu)$ AND $V_s(a, b; \nu)$

The values of $M_s(a, b)$, $W_s(a, b: \nu)$ and $V_s(a, b: \nu)$ may be generated using

$$\frac{a+b+s+1}{s} M_{s+1}(a, b) = \frac{b-a}{b+a} M_s(a, b) - \frac{ab}{(a+b)^2} M_{s-1}(a, b), \tag{6·1}$$

$$V_{s+1}(a, b: \nu) + \nu V_s(a, b: \nu) = W_s(a, b: \nu), \tag{6·2}$$

$$(a+b+s) W_{s+1}(a, b: \nu) = a - \nu(a+b) + s(1-2\nu) W_s(a, b: \nu) + s\nu(1-\nu) W_{s-1}(a, b: \nu), \tag{6·3}$$

where we will define

$$\left.\begin{aligned} sW_{s-1}(a, b: \nu) &= -\beta_\nu(a, b), && s = 0 \\ &= \int_0^\nu \beta_P(a, b)\, dP, && s = 1 \end{aligned}\right\}$$

and

$$V_0(a, b: \nu) = \frac{a+b-1}{a-1} \int_0^\nu \beta_P(a-1, b)\, dP.$$

Equation (6·1) is a result which follows directly from the differential equation which underlies the Pearson system of frequency curves. Equation (6·3) follows using a similar technique, whilst equation (6·2) is obtained by simply combining the two integrals implicit in the left-hand side of the equation.

7. PLACKETT'S METHOD

The essential details of Plackett's method are reproduced here for purposes of comparison.

Plackett defines

$$L(x) = \log \frac{P(x)}{1-P(x)}. \tag{7·1}$$

Then $K_j(r, n)$, the jth cumulant of $L(x_{r:n})$, is given by

$$K_j(r, n) = D_w^j \{\log \Gamma(r+w)\, \Gamma(n-r+1-w)\}|_{w=0}. \tag{7·2}$$

These cumulants are either tabulated in terms of the polygamma functions or may be expressed as a finite series.

Using a Taylor series expansion for $x_{r:n}$ about the point where $L(x_{r:n}) = K_1(r, n)$ we obtain after integrating over the p.d.f. of $x_{r:n}$

$$\mathscr{E}(x_{r:n}) = \sum_{j=0}^{s} \frac{1}{j!} x_j(L = K_1)\, \mu_j(r, n) + \frac{1}{(s+1)!} \mathscr{E}\{(L-K_1)^{s+1} x_{s+1}[L = K_1 + \gamma(L-K_1)]\}, \tag{7·3}$$

where

$$\mu_j(r, n) = \mathscr{E}[L(x_{r:n}) - K_1]^j \quad \text{and} \quad x_j(L) = \left(\frac{d}{dL}\right)^j x(L). \tag{7·4}$$

Derivatives and moments in the finite sum may be evaluated. Plackett notes that $x_j(L)$ is bounded and provides the table

j	1	2	3	4
$\max_L (x_j(L))$	0·62666	0·07376	0·06724	0·04597

$$\mathscr{E}|L-K_1|^{s+1} = \mathscr{E}(L-K_1)^{s+1} \qquad (s \text{ odd}),$$

also

$$\mathscr{E}|L-K_1|^{s+1} < \{\mathscr{E}(L-K_1)^{s+2}\}^{(s+1)/(s+2)} \qquad (s \text{ even}),$$

so that we have finally

$$\left|\mathscr{E}(x_{r:n}) - \sum_{j=0}^{s} x_j(L = K_1)\,\mu_j(r,n)\right| < \max_L (x_{s+1}(L))\,\lambda_{s+1}(r,n), \tag{7·5}$$

with
$$\lambda_j(r,n) = \mu_j(r,n) \quad (j \text{ even}),$$
$$= (\mu_{j+1}(r,n))^{j/(j+1)} \quad (j \text{ odd}).$$

8. A COMPARISON OF THE TWO METHODS

To illustrate the methods, we consider the case when $r = 6$; $n = 9$. The true value of $\mathscr{E}(x_{6:9})$ is 0·2745 2592 to eight places of decimals (see Teichroew, 1956). Thus we are able to give the true error after a given number of terms of either series and compare this true error with the bounds for error obtained from equations (5·11), (5·12) and (7·5).

David and Johnson series after 2m terms

		Bounds for error	
m	True error	Lower	Upper
1	+0·002662	−0·002372	+0·010997
2	+0·000668	−0·001176	+0·003668
3	+0·000127	−0·000729	+0·001117
4	+0·000043	−0·000541	+0·000424

Plackett series after s terms

		Absolute bound
s	True error	for error
2	+0·001240	0·009078
3	+0·001714	0·001753
4	+0·000580	—

Term for term (i.e. setting $s = 2m$) the Plackett series seems to converge a little more rapidly than does the David & Johnson series. Plackett has not given $\max(x_j(L))$ above $j = 4$, so that the bound for error after 4 terms is not given in the table above (these maxima are very laborious to locate although once obtained they are relevant to any r and n). It would seem that the Plackett bounds are a little finer than those which were obtained for the David & Johnson series.

The slight superiority of the Plackett series in respect of convergence is overshadowed by the computational advantages of the David & Johnson technique. If we return to equation (3·4) we note that the derivatives $x_j(p_r)$ are functions of p_r only and are therefore constant with constant p_r. In addition, the moments $M_j(r, n-r+1)$ may be expressed as a power series in $1/(n+2)$ with coefficients which are functions of p_r only. For example

$$M_2(r, n-r+1) = p_r q_r (n+2)^{-1},$$
$$M_3(r, n-r+1) = 2(q_r - p_r)\,p_r q_r \sum_{i=2}^{\infty} (n+2)^{-i}.$$

This enables us to write
$$\mathscr{E}(x_{r:n}) = \sum_{i=0}^{\infty} H_i(p_r, 0, 1)\,(n+2)^{-i},$$

where $H_i(p_r, 0, 1)$ is a numerical coefficient which varies with $p_r = r/(n+1)$ only. Once these coefficients have been obtained for a range of values of p_r, it is a very simple matter to

calculate $\mathscr{E}(x_{r:n})$ for any values of r and n. David & Johnson have given tables of the algebraic forms of the coefficients which occur in the expansion of $K(x_{i:n}^a x_{j:n}^b x_{k:n}^c x_{s:n}^d)$ in terms of a power series in $1/(n+2)$, for $a+b+c+d \leqslant 4$, as far as the coefficient in $(n+2)^{-3}$.

The author (Saw, 1958) has given the numerical values of the coefficients $H_i(p_r, a, b)$ which occur in the expansion of the integral

$$\int_0^1 \left\{\frac{Z[x(P)]}{P}\right\}^a \{x(P)\}^b \beta_P(r, n-r+1)\, dP = \sum_{i=0}^{\infty} H_i(p_r, a, b)\,(n+2)^{-i},$$

for $p_r = 0{\cdot}50\,(0{\cdot}05)\,0{\cdot}80$; $i = 0\,(1)\,5$; $a+b \leqslant 4$.

On the other hand, the coefficients in the Plackett series are functions of $K_1(r, n)$. Now using equation (7·2)

$$K_1(r, n) = \frac{1}{n-r+1} + \frac{1}{n-r+2} + \ldots + \frac{1}{r-1},$$

which does not remain constant as some function of r and n remains constant. Hence we need to calculate a new Plackett series for every combination of r and n.

I wish to record my gratitude to Dr F. N. David, Dr D. E. Barton and Dr C. L. Mallows for their encouragement in carrying out this investigation.

REFERENCES

DAVID, F. N. & JOHNSON, N. L. (1954). Statistical treatment of censored data. *Biometrika*, **41**, 228–40.

PLACKETT, R. L. (1958). Linear estimation from censored data. *Ann. Math. Statist.* **29**, 131–42.

SAW, J. G. (1958). Moments of sample moments of censored samples from a normal population. *Biometrika*, **45**, 211–21.

TEICHROEW, D. (1956). Tables of expected values of ordered statistics and products of ordered statistics for samples of size twenty and less from a normal population. *Ann. Math. Statist.* **27**, 410–26.

Biometrika (1960), **47**, 1 *and* 2, *p.* 87
Printed in Great Britain

More significance tests on the sphere

BY G. S. WATSON†
University of Toronto

1. INTRODUCTION

It has been plausibly suggested (Fisher, 1953; Watson & Irving, 1957) that the directions of remanent magnetization of specimens from the same source of suitable rock are distributed on the unit sphere according to the formula

$$\frac{\kappa}{4\pi \sinh \kappa} \exp(\kappa \cos\theta). \tag{1·1}$$

Here $\cos\theta = \lambda l + \mu m + \nu n$, (l, m, n) are the direction cosines of an observed direction, (λ, μ, ν) are the direction cosines of the mean direction, and κ is an accuracy parameter, large values of which lead to small dispersions. This paper is primarily concerned with a question which is sometimes of geophysical importance: are the mean directions of a set of populations all in the same plane?

The basic distribution theory of (1·1) was given by Fisher (1953). For a sample of N from (1·1), the maximum likelihood (M.L.) estimator or (λ, μ, ν) is the set of direction cosines, (l, m, n) say, of the vector resultant of the sample of unit vectors. If the length of the resultant is denoted by R, then the M.L. estimator of κ, when (λ, μ, ν) is also estimated, is the solution of

$$\coth k - \frac{1}{k} = \frac{R}{N}. \tag{1·2}$$

When N is large and R is near N, this has the approximate solution

$$k = \frac{N-1}{N-R}. \tag{1·3}$$

When (λ, μ, ν) is known, the M.L. equation for κ is the same as (1·2) with R replaced by $X = R\cos\psi$, where ψ is the angle between the resultant and (λ, μ, ν). The approximate estimator k_0 of κ is then given by

$$k_0 = \frac{N}{N-X}. \tag{1·4}$$

It may be shown (Watson, 1956) that, for large κ,

$$\left.\begin{aligned} 2\kappa(N-X) &\approx \chi^2_{2N}, \\ 2\kappa(N-R) &\approx \chi^2_{2(N-1)}. \end{aligned}\right\} \tag{1·5}$$

These results and a comparison of the identity

$$\Sigma(x_i-\mu)^2 = \Sigma(x_i-\bar{x})^2 + N(\bar{x}-\mu)^2,$$

with

$$2\kappa(N-X) = 2\kappa(N-R) + 2\kappa(R-X), \tag{1·6}$$

leads to the suggestion that‡

$$(N-1)\frac{(R-X)}{N-R} \approx F_{2,2(N-1)}, \tag{1·7}$$

† This work was carried out while the author was visiting Princeton University and prepared in connexion with research sponsored by the Office of Naval Research.

be used as a test of a prescribed mean direction. Further tests were derived by the same argument in Watson (1956). The same approach yields the procedures for analysing the coplanarity of mean vectors that will be given below. However, it is of interest to deduce them by likelihood-ratio methods. It will be apparent then that the above tests may be derived by this method too.

The generalization of the foregoing to unit vectors in any number of dimensions was indicated in Watson & Williams (1956). The same is easily possible here. A brief review of studies of the dimensionality of the mean vectors of multivariate normal populations is given by Anderson (1958). For large samples a satisfactory solution based on chi-square is possible but, in small samples, difficult problems are posed by nuisance parameters. The distribution (1·1) presents the same difficulties.

2. A TEST FOR COPLANARITY

Consider p populations (1·1) with mean vectors $(\lambda_{(i)}, \mu_{(i)}, \nu_{(i)})$ $(i = 1, \ldots, p)$ and all with the same accuracy parameter κ. Suppose that, for $i = 1, \ldots, p$, a sample of N_i has been drawn from the ith population and that it yields a resultant with direction cosines (l_i, m_i, n_i) and of length R_i. Writing $\sum_{i=1}^{p} N_i = N$, the logarithm of the likelihood of the entire set of observations is proportional to

$$N(\log \kappa - \log \sinh \kappa) + \kappa \sum_{i=1}^{p} R_i(l_i\lambda_{(i)} + m_i\mu_{(i)} + n_i\nu_{(i)}). \tag{2·1}$$

If a known direction (λ, μ, ν) is orthogonal to all the population means, it is easily shown that the M.L. estimator of $(\lambda_{(i)}, \mu_{(i)}, \nu_{(i)})$ is given by

$$(\hat{\lambda}_{(i)}, \hat{\mu}_{(i)}, \hat{\nu}_{(i)}) = \frac{[l_i - (\Sigma\lambda l_i)\lambda, m_i - (\Sigma\lambda l_i)\mu, n_i - (\Sigma\lambda l_i)\nu]}{\sqrt{[1 - (\Sigma\lambda l_i)^2]}}, \tag{2·2}$$

where $\Sigma\lambda l_i = \lambda l_i + \mu m_i + \nu n_i$. Under the same conditions, the estimator of κ is given by the solution of

$$\coth \hat{\kappa} - \frac{1}{\hat{\kappa}} = \frac{\sum_{i=1}^{p} R_i \sqrt{[1 - (\Sigma\lambda l_i)^2]}}{N}. \tag{2·3}$$

A comparison with (1·2) shows that $\Sigma R_i \sqrt{[1 - (\Sigma\lambda l_i)^2]}$ takes the place of R. If no restriction is placed on the mean vectors, then (l_i, m_i, n_i) is the M.L. estimator of $(\lambda_{(i)}, \mu_{(i)}, \nu_{(i)})$ and the estimator of κ is the solution of

$$\coth \kappa^* - \frac{1}{\kappa^*} = \frac{\sum_{i=1}^{p} R_i}{N}. \tag{2·4}$$

To test the null hypothesis that all the mean vectors lie in a plane normal to the prescribed vector (λ, μ, ν), we form the logarithm of the likelihood ratio in the usual way. On the null hypothesis, (2·1) becomes

$$N(\log \hat{\kappa} - \log \sinh \hat{\kappa}) + \hat{\kappa} \sum_{i=1}^{p} R_i \sqrt{[1 - (\Sigma\lambda l_i)^2]}, \tag{2·5}$$

and since we are going to assume that N and κ are large, (2·5) reduces to

$$N(\log\hat{\kappa}-1), \tag{2·6}$$

where

$$\hat{\kappa}=\frac{N}{N-\Sigma R_i\sqrt{[1-(\Sigma\lambda l_i)^2]}}. \tag{2·7}$$

On the alternative hypothesis (2·1) again takes on the form (2·6) but with κ^* instead of $\hat{\kappa}$, where

$$\kappa^*=\frac{N}{N-\Sigma R_i}. \tag{2·8}$$

Thus, by Wilks' Theorem,

$$-2N\log(\hat{\kappa}/\kappa^*)\sim\chi^2_p. \tag{2·9}$$

Then (2·9) may be written as

$$2N\log\left\{1+\frac{\sum_{i=1}^{p}R_i(1-\sqrt{[1-(\Sigma\lambda l_i)^2]})}{\Sigma(N_i-R_i)}\right\}\sim\chi^2_p.$$

When all the N_i are large, this may further be approximated by

$$2N\frac{\sum_{i=1}^{p}R_i(1-\sqrt{[1-(\Sigma\lambda l_i)^2]})}{\sum_{i=1}^{p}(N_i-R_i)}\sim\chi^2_p, \tag{2·10}$$

or finally as

$$N\frac{\boldsymbol{\lambda}'\mathbf{U}\boldsymbol{\lambda}}{\Sigma(N_i-R_i)}\sim\chi^2_p, \tag{2·11}$$

where $\boldsymbol{\lambda}'=(\lambda,\mu,\nu)$, and

$$\mathbf{U}=\begin{Bmatrix}\Sigma R_i l_i^2, & \Sigma R_i l_i m_i, & \Sigma R_i l_i n_i,\\ \Sigma R_i m_i l_i, & \Sigma R_i m_i^2, & \Sigma R_i m_i n_i,\\ \Sigma R_i n_i l_i, & \Sigma R_i n_i m_i, & \Sigma R_i n_i^2.\end{Bmatrix} \tag{2·12}$$

Had the analysis of variance analogue been pursued, it would have suggested that

$$\frac{N-p}{p}\frac{\boldsymbol{\lambda}'\mathbf{U}\boldsymbol{\lambda}}{\Sigma(N_i-R_i)}\approx F_{p,2(N-p)}, \tag{2·13}$$

which for large N is the same as (2·11). Thus (2·11) or (2·13) provide two approximate tests of the hypothesis that a set of p mean vectors lie in a plane normal to $\boldsymbol{\lambda}$.

The same likelihood-ratio procedure gives, as a test of coplanarity only, the statistic

$$N\frac{\min_{\boldsymbol{\lambda}}\boldsymbol{\lambda}'\mathbf{U}\boldsymbol{\lambda}}{\Sigma(N_i-R_i)}\sim\chi^2_{p-2}. \tag{2·14}$$

The analysis of variance analogy gives

$$\frac{N-p}{p-2}\frac{\min_{\boldsymbol{\lambda}}\boldsymbol{\lambda}'\mathbf{U}\boldsymbol{\lambda}}{\Sigma(N_i-R_i)}\approx F_{p-2,2(N-p)}. \tag{2·15}$$

In this case the value of $\boldsymbol{\lambda}$ minimizing $\boldsymbol{\lambda}'\mathbf{U}\boldsymbol{\lambda}$ is clearly the latent vector of $\mathbf{U}$ corresponding to its least latent root and this vector is the M.L. estimator of the normal to the best fitting plane. The tests (2·13) and (2·15) can be conveniently arranged in an analysis of variance table, as

is shown in our example. It is natural in practice to test first for coplanarity and then to test the adequacy of a prescribed normal. This latter test is either of

$$N\frac{[\boldsymbol{\lambda}'\mathbf{U}\boldsymbol{\lambda}-\min_{\boldsymbol{\lambda}}\boldsymbol{\lambda}'\mathbf{U}\boldsymbol{\lambda}]}{\Sigma(N_i-R_i)}\sim\chi^2_2, \tag{2·16}$$

or

$$\frac{(N-p)}{2}\frac{\boldsymbol{\lambda}'\mathbf{U}\boldsymbol{\lambda}-\min_{\boldsymbol{\lambda}}\boldsymbol{\lambda}'\mathbf{U}\boldsymbol{\lambda}}{\Sigma(N_i-R_i)}\approx F_{2,2(N-p)}. \tag{2·17}$$

These tests provide a confidence cone for $\boldsymbol{\lambda}$—the set of all non-significant $\boldsymbol{\lambda}$. This will be illustrated in our example.

A procedure which should be more robust against differences between the accuracies in the populations is provided by

$$2\sum_{i=1}^{p}\frac{N_i-1}{N_i-R_i}R_i(1-\sqrt{[1-(\Sigma\lambda l_i)^2]})\approx\chi^2_p. \tag{2·18}$$

Writing

$$\mathbf{W}=\begin{bmatrix}\Sigma\frac{R_i(N_i-1)}{N_i-R_i}l_i^2, & \Sigma\frac{R_i(N_i-1)}{N_i-R_i}l_im_i, & \Sigma\frac{R_i(N_i-1)}{N_i-R_i}l_in_i,\\ \Sigma\frac{R_i(N_i-1)}{N_i-R_i}m_il_i, & \Sigma\frac{R_i(N_i-1)}{N_i-R_i}m_i^2, & \Sigma\frac{R_i(N_i-1)}{N_i-R_i}m_in_i,\\ \Sigma\frac{R_i(N_i-1)}{N_i-R_i}n_il_i, & \Sigma\frac{R_i(N_i-1)}{N_i-R_i}n_im_i, & \Sigma\frac{R_i(N_i-1)}{N_i-R_i}n_i^2,\end{bmatrix} \tag{2·19}$$

this leads to an approximate analysis,

$$\left.\begin{aligned}\min_{\boldsymbol{\lambda}}\boldsymbol{\lambda}'\mathbf{W}\boldsymbol{\lambda}&=\text{'Devia'ons from coplanarity'}\approx\chi^2_{p-2},\\ \boldsymbol{\lambda}'\mathbf{W}\boldsymbol{\lambda}-\min_{\boldsymbol{\lambda}}\boldsymbol{\lambda}'\mathbf{W}\boldsymbol{\lambda}&=\text{'Deviations from assumed normal'}\approx\chi^2_2.\end{aligned}\right\} \tag{2·20}$$

3. An example

A part of a larger collection of data (kindly supplied by Dr K. M. Creer) whose analysis and geophysical significance will be discussed elsewhere, may be used to illustrate the foregoing procedures.

Samples were taken from three populations whose mean vectors were, on a certain supposition, coplanar. The table below gives the sample sizes, the direction cosines and lengths of the sample resultants and the within-sample estimates of κ. These figures contain all the relevant information in the data.

Population	l	m	n	R	N	k
A	−0·0698	−0·9589	−0·2750	33·172	35	18·6
B	−0·8572	0·2575	0·4460	8·567	9	18·5
C	−0·5469	−0·7302	−0·4095	5·786	6	23·4

The matrices U and W and their latent roots and vectors were computed electronically by Dr Creer. The least roots of U and W are respectively 0·02682 and 0·5900. Hence

$$\frac{2(N-p)}{p-2}\frac{\frac{1}{2}\min\boldsymbol{\lambda}'\mathbf{U}\boldsymbol{\lambda}}{\Sigma(N_i-R_i)}=47\times\frac{0\cdot02682}{2\cdot475}=0\cdot509$$

is to be referred to the F-distribution with 1 and 94 D.F. The other root 0·590 is to be referred

to the chi-square distribution with 1 D.F. The close agreement of the two statistics is a reflexion of the very similar values of k in the three populations. There is no evidence of significant non-coplanarity.

To put a cone of confidence around the normal to this common plane, it is natural to use the latent vectors as new co-ordinate axes. In this case (2·15) may be written, for 95 % confidence,

$$\tfrac{1}{2}(47)\,(9{\cdot}9497\lambda^{*2} + 37{\cdot}5481\mu^{*2} + 0{\cdot}02682\nu^{*2} - 0{\cdot}02682) \leqslant 3{\cdot}05.$$

Writing $\nu^* = \cos\theta$, $\lambda^* = \sin\theta\cos\phi$, $\mu^* = \sin\theta\sin\phi$, this becomes

$$\sin^2\theta(9{\cdot}9497\cos^2\phi + 37{\cdot}5481\sin^2\phi) + 0{\cdot}02682\cos^2\theta \leqslant 0{\cdot}103.$$

The cross-section of this cone is elliptical. The average semi-angle $\bar{\theta}$ is given by

$$23{\cdot}7489\sin^2\bar{\theta} + 0{\cdot}02682\cos^2\bar{\theta} = 0{\cdot}103$$

so that $\bar{\theta} = 3°\,15'$. However, $\theta_{\text{max.}}$, which occurs when $\phi = 0°, 180°$, is 5° while $\theta_{\text{min.}}$, which occurs for $\phi = 90°, 270°$, is 2° 35′. Since the latent vector of U corresponding to the minimum root is

$$(0{\cdot}4802, 0{\cdot}2004, -0{\cdot}8540),$$

the 95 % cone of confidence for the normal to the plane in which the three population means lie has this latent vector as axis and a semi-angle between $2\frac{1}{2}°$ and 5°, with a mean of 3°.

The cone of confidence obtained from the weighted analysis using **W** is very similar. For example, the cosine of the angle between the axes of the two cones is 0·9999927 so that the angle is less than 6′.

References

Anderson, T. W. (1958). *An Introduction to Multivariate Analysis.* New York: John Wiley and Sons.

Fisher, R. A. (1953). Dispersion on a sphere. *Proc. Roy. Soc.* A, **217**, 295.

Watson, G. S. (1956). Analysis of dispersion on a sphere. *Mon. Not. R. Astr. Soc. Geophys. Suppl.* **7**, 153.

Watson, G. S. & Irving, E. (1957). Statistical methods in rock magnetism. *Mon. Not. R. Astr. Soc. Geophys. Suppl.* **7**, 289.

Watson, G. S. & Williams, E. J. (1956). Significance tests on the circle and the sphere. *Biometrika*, **43**, 344.

Biometrika (1960), **47**, 1 *and* 2, *p.* 93
Printed in Great Britain

An approximation to the multinomial distribution: some properties and applications

BY N. L. JOHNSON
University College London

1. When investigating the properties of statistics constructed from discontinuous random variables, it is often useful to have a reasonably good approximation in terms of continuous variables. The approximation to the distribution of the contingency table χ^2 statistic by means of a continuous Pearson Type III variable is a well-known example. In searching for such an approximate representation it is a common practice to try to arrange that certain important properties of the original variables—range of variation, and/or lower order moments—shall be preserved. Most approximations of this type have been worked out for single variables, but there is no reason, in principle, why similar approximations cannot be used for multivariate systems.

In this paper the properties of an approximation to the multinomial distribution

$$P(n_1, n_2, \ldots, n_k) = N! \prod_{j=1}^{k} (\pi_j^{n_j}/n_j!) \quad \left(\sum_1^k \pi_j = 1;\ \sum_1^k n_j = N\right) \tag{1}$$

will be discussed in some detail. The approximation is obtained by regarding the relative frequencies $f_j = n_j/N$ as continuous variables with joint probability density functions derived from

$$g(f_1, f_2, \ldots, f_k) = \Gamma(N-1) \prod_{j=1}^{k} [f_j^{(N-1)\pi_j - 1}/\Gamma((N-1)\pi_j)]. \tag{2}$$

The joint probability density function of any $k-1$ out of the k f's is obtained by replacing the missing f_l by $1 - \sum_{j \neq l} f_j$ in $g(f_1, f_2, \ldots, f_k)$.

This approximation gives the correct first and second moments and product moments for the joint distribution of the f's, and also the correct range of variation. It has been used by Johnson & Young (1960) to derive approximate distributions for the range, $(\max n_j - \min n_j)$, and for $\max n_j$. Some other applications are discussed in the present paper.

2. From (2) it follows that the general product moment about zero,

$$\mu'_{\alpha}(f) = \mu'_{\alpha_1, \alpha_2, \ldots, \alpha_k}(f) = \mathscr{E}\left(\prod_{j=1}^{k} f_j^{\alpha_j}\right)$$

is equal to

$$\frac{\Gamma(N-1)}{\Gamma\left(N-1+\sum_1^k \alpha_j\right)} \prod_{j=1}^{k} \frac{\Gamma((N-1)\pi_j + \alpha_j)}{\Gamma((N-1)\pi_j)} \tag{3}$$

or, equivalently,

$$\mu'_{\alpha}(n) = \mu'_{\alpha_1, \alpha_2, \ldots, \alpha_k}(n) = \mathscr{E}\left(\prod_{j=1}^{k} n_j^{\alpha_j}\right) = \frac{N^{\Sigma\alpha_j}\Gamma(N-1)}{\Gamma\left(N-1+\sum_1^k \alpha_j\right)} \prod_{j=1}^{k} \frac{\Gamma((N-1)\pi_j + \alpha_j)}{\Gamma((N-1)\pi_j)}.$$

The exact multinomial factorial moments of the n_j's, based on (1), are

$$\mu_{(\theta)}(n) = \mu_{(\theta_1, \theta_2, \ldots, \theta_k)}(n) = \mathscr{E}\left(\prod_{j=1}^{k} n_j^{(\theta_j)}\right) = N^{(\Sigma\theta_j)} \prod_{j=1}^{k} \pi_j^{\theta_j}. \tag{4}$$

Considering now the first- and second-order moments and product moments, both (3) and (4) give $\mathscr{E}(f_j) = \pi_j$; $\operatorname{var}(f_j) = \pi_j(1-\pi_j)/N$; $\operatorname{cov}(f_i, f_j) = -\pi_i\pi_j/N$.

This agreement between approximation (2) and the multinomial distribution (1) does not extend to the higher moments. The table below compares the values of $\sqrt{\beta_1}$ and β_2 of f_j for the two distributions

	Exact (1)
$\sqrt{\beta_1}$	$(1-2\pi_j)/[N\pi_j(1-\pi_j)]^{\frac{1}{2}}$
β_2	$3+[1-6\pi_j(1-\pi_j)]/[N\pi_j(1-\pi_j)]$

	Approximation (2)
$\sqrt{\beta_1}$	$[2N/(N+1)]\,(1-2\pi_j)/[N\pi_j(1-\pi_j)]^{\frac{1}{2}}$
β_2	$3N(N-7)/(N+1)\,(N+2)+[6N^2/(N+1)\,(N+2)]\,[1-\pi_j(1-\pi_j)]/[N\pi_j(1-\pi_j)]$

The approximation gives skewness twice as large as the exact value, and kurtosis nearly six times as large. As we shall see in the next section, however, these disagreements are not reflected in any serious discrepancy in calculated probabilities.

We may note that the exact value for $\mathscr{E}(f_i^2 f_j)$ is

$$(N-1)\,\pi_i\pi_j[(N-2)\,\pi_i+1]\,N^{-2},$$

while (2) gives $$(N-1)\,\pi_i\pi_j[(N-1)\,\pi_i+1]\,N^{-1}(N+1)^{-1}.$$

The expected value of f_j^{-1} given by (3) is

$$(N-2)/[(N-1)\,\pi_j-1].$$

In fact, f_j^{-1} has an infinite expected value since $\Pr\{f_j = 0\} > 0$. However, if the distribution of f_j be truncated by omission of the value $f_j = 0$ the expected value is finite. Such distributions (or, equivalently, those of n_j) have been studied by Fieller & Hartley (1954) and David & Johnson (1956) *inter alia*. The latter authors obtained an expansion of the form

$$\mathscr{E}(f_j^{-1}) = \frac{1}{\pi_j}\left[1+\frac{1-\pi_j}{N\pi_j}+\frac{(2-\pi_j)\,(1-\pi_j)}{(N\pi_j)^2}+\ldots\right].$$

Expanding the value $(N-2)/[(N-1)\,\pi_j-1]$ derived from approximation (2) in powers of N^{-1} we get

$$\mathscr{E}(f_j^{-1}) \doteq \frac{1}{\pi_j}\left[1+\frac{1-\pi_j}{N\pi_j}+\frac{(1+\pi_j)\,(1-\pi_j)}{(N\pi_j)^2}+\ldots\right].$$

The two expressions (up to terms in N^{-2}) agree to order N^{-1} and differ only by

$$(1-2\pi_j)\,(1-\pi_j)/N^2\pi_j^3.$$

3. From (2), the approximate joint probability density function of $f_1, f_2, \ldots, f_{k-1}$ is

$$p(f_1, f_2, \ldots, f_{k-1}) = \frac{\Gamma(N-1)}{\Gamma((N-1)\,\pi_k)}\left(1-\sum_{j=1}^{k-1} f_j\right)^{(N-1)\pi_k-1}\prod_{j=1}^{k-1}\frac{f_j^{(N-1)\pi_j-1}}{\Gamma((N-1)\,\pi_j)} \quad \left(\sum_{j=1}^{k-1} f_j \leqslant 1,\ 0 \leqslant f_j\right). \tag{5a}$$

Integrating out f_{k-1} the approximate joint probability density function of $f_1, f_2, \ldots, f_{k-2}$ is

$$p(f_1, f_2, \ldots, f_{k-2}) = \frac{\Gamma(N-1)}{\Gamma((N-1)\,(\pi_{k-1}+\pi_k))}\left(1-\sum_{j=1}^{k-2} f_j\right)^{(N-1)(\pi_{k-1}+\pi_k)-1}\prod_{j=1}^{k-2}\frac{f_j^{(N-1)\pi_j-1}}{\Gamma((N-1)\,\pi_j)}$$
$$\left(\sum_{j=1}^{k-2} f_j \leqslant 1,\ 0 \leqslant f_j\right). \tag{5b}$$

Repeating the process we find

$$p(f_1, f_2, \ldots, f_s) = \frac{\Gamma(N-1)}{\Gamma\left((N-1)\sum_{s+1}^{k}\pi_j\right)}\left(1-\sum_{j=1}^{s}f_j\right)^{(N-1)\sum_{s+1}^{k}\pi_j-1}\prod_{j=1}^{s}\frac{f_j^{(N-1)\pi_j-1}}{\Gamma((N-1)\pi_j)}\quad\left(\sum_{j=1}^{s}f_j\leqslant 1, f_j\geqslant 0\right), \tag{5c}$$

and $$p(f_j) = \frac{\Gamma(N-1)}{\Gamma((N-1)\pi_j)\,\Gamma((N-1)(1-\pi_j))}f_j^{(N-1)\pi_j-1}(1-f_j)^{(N-1)(1-\pi_j)-1}\quad(0\leqslant f_j\leqslant 1). \tag{5d}$$

We see that the approximate distributions of subsets of the f's, derived from (2), are of the same form as the initial function $g(f_1, f_2, \ldots, f_k)$. The approximation is thus self-consistent in this respect.

Table 1. *Type I approximation to the binomial*

$x = I_{(r-\frac{1}{2})/N}((N-1)\pi, (N-1)(1-\pi))$ (approximate values) $\qquad y = I_{1-\pi}(N-r+1, r)$ (exact values)

	$\pi = 0{\cdot}1$		$\pi = 0{\cdot}3$		$\pi = 0{\cdot}5$	
r	x	y	x	y	x	y
			$N = 25$			
1	0·0359	0·0717	—	0·0001	—	—
3	0·5728	0·5371	0·0043	0·0090	—	—
5	0·8959	0·9020	0·0880	0·0905	0·0002	0·0002
7	0·9828	0·9905	0·3558	0·3407	0·0065	0·0073
8	0·9936	0·9977	0·5239	0·5118	0·0214	0·0216
10	0·9994	1·0000	0·8083	0·8105	0·1188	0·1147
12	1·0000	—	0·9449	0·9557	0·3488	0·3450
15	—	—	0·9974	0·9982	0·7825	0·7878
18	—	—	1·0000	1·0000	0·9786	0·9784
20	—	—	—	—	0·9985	0·9980

	$\pi = 0{\cdot}1$		$\pi = 0{\cdot}3$		$\pi = 0{\cdot}5$	
r	x	y	x	y	x	y
			$N = 50$			
1	0·0001	0·0051	—	—	—	—
3	0·0999	0·1117	—	—	—	—
5	0·4559	0·4312	—	0·0002	—	—
7	0·7791	0·7702	0·0012	0·0024	—	—
10	0·9676	0·9754	0·0361	0·0402	—	—
12	0·9933	0·9968	0·1399	0·1390	—	—
15	0·9996	0·9999	0·4574	0·4468	0·0011	0·0013
17	1·0000	1·0000	0·6878	0·6838	0·0073	0·0076
20	—	—	0·9118	0·9151	0·0604	0·0595
22	—	—	0·9716	0·9749	0·1635	0·1611
25	—	—	0·9968	0·9976	0·4446	0·4439
27	—	—	0·9995	0·9997	0·6622	0·6641
30	—	—	1·0000	1·0000	0·8970	0·8987
32	—	—	—	—	0·9675	0·9676
35	—	—	—	—	0·9970	0·9967
37	—	—	—	—	0·9997	0·9996
40	—	—	—	—	1·0000	1·0000

From (5*d*) $\Pr\{n_j < r_j\} \doteq I_{(r_j-\frac{1}{2})/N}((N-1)\pi_j, (N-1)(1-\pi_j))$ $(r_j > 0)$, (6*a*)

where $$I_X(m,n) = \frac{1}{B(m,n)}\int_0^X x^{m-1}(1-x)^{n-1}\,dx$$

is the incomplete beta-function ratio. The exact value of this probability is

$$\Pr\{n_j < r_j\} = I_{1-\pi_j}(N-r_j+1, r_j). \tag{6b}$$

The work of Davies (1933) on approximating the hypergeometric distribution $F(\alpha,\beta;\gamma;1)$ by a Type I curve indicated that Type I should also give a good approximation to a binomial. This was confirmed by Benedetti (1956).

Approximate and exact values calculated from (6*a*) and (6*b*), respectively, are compared in Table 1. Benedetti (1956) gives corresponding values for $N = 100$. The comparison indicates that approximation (2) is reasonably good, over the bulk of the probability mass, in regard to the marginal distributions it implies.

We have already seen, in § 2, that the approximation is not exact for moments of order higher than two (or less than zero). In § 7 we will also show that it is not self-consistent in regard to conditional distributions (of one subset of f's, given another subset).

4. From equations (5) we can obtain the distributions of sums of subsets of the f_j's by the usual processes of probability calculus. We find, for instance,

$$p\left(\sum_{j=1}^{s} f_j\right) = \frac{\Gamma(N-1)}{\Gamma\left((N-1)\sum_1^s \pi_j\right)\Gamma\left((N-1)\sum_{s+1}^{k}\pi_j\right)}\left(\sum_1^s f_j\right)^{(N-1)\sum_1^s \pi_j - 1}\left(1-\sum_1^s f_j\right)^{(N-1)\sum_{s+1}^{k}\pi_j - 1}$$

$$\left(0 \leqslant f_j;\ \sum_{j=1}^{s} f_j \leqslant 1\right). \tag{5e}$$

This is of the same form as $p(f_j)$ in (5*d*) with π_j replaced by $\sum_1^s \pi_j$. Since

$$\sum_1^s f_j = \left(\sum_1^s n_j\right)\Big/N$$

is the proportion of observations falling in the superclass formed by combining the first s classes of the original multinomial, the approximation is self-consistent in this respect also.

Equation (5*e*) (and similar equations) may also be obtained by the following instructive analogy.

Let $v_1, v_2, \ldots, v_k$ be mutually independent random variables distributed as χ^2 with $\nu_1, \nu_2, \ldots, \nu_k$ degrees of freedom, respectively. The joint probability density function of the v's is then

$$p(v_1, v_2, \ldots, v_k) = \prod_{j=1}^{k}\left[2^{-\frac{1}{2}\nu_j}\{\Gamma(\tfrac{1}{2}\nu_j)\}^{-1} v_j^{\frac{1}{2}\nu_j - 1} e^{-\frac{1}{2}v_j}\right] \quad (0 \leqslant v_j).$$

Now apply the transformation

$$y_i = v_i\Big/\sum_{j=1}^{k} v_j \quad (i = 1, 2, \ldots, k-1); \quad y_k = \sum_{j=1}^{k} v_j,$$

so that $$v_i = y_i y_k \quad (i = 1, 2, \ldots, k-1); \quad v_k = y_k\left(1 - \sum_{j=1}^{k-1} y_j\right).$$

The Jacobian of this transformation is

$$\frac{\partial(v_1, v_2, \ldots, v_k)}{\partial(y_1, y_2, \ldots, y_k)} = y_k^{k-1}\left(1 - \sum_{j=1}^{k-1} y_j\right)$$

and so the joint probability density function of $y_1, y_2, \ldots, y_k$ is

$$p(y_1, y_2, \ldots, y_k) = \frac{y_k^{\frac{1}{2}\sum_1^k \nu_j - 1} e^{-\frac{1}{2}y_k}}{2^{\frac{1}{2}\sum_1^k \nu_j}\Gamma\left(\frac{1}{2}\sum_1^k \nu_j\right)}\left[\frac{\Gamma\left(\frac{1}{2}\sum_1^k \nu_j\right)}{\prod_1^k \Gamma(\frac{1}{2}\nu_j)}\left(1 - \sum_1^{k-1} y_j\right)^{\frac{1}{2}\nu_k - 1}\prod_{j=1}^{k-1} y_j^{\frac{1}{2}\nu_j - 1}\right]$$

$$\left(0 \leqslant y_j;\ \sum_1^{k-1} y_j \leqslant 1\right).$$

Integrating out y_k, we find the joint probability density function of the $(k-1)$ ratios $y_1, y_2, \ldots, y_{k-1}$

$$p(y_1, y_2, \ldots, y_{k-1}) = \frac{\Gamma\left(\frac{1}{2}\sum_1^k \nu_j\right)}{\prod_1^k \Gamma(\frac{1}{2}\nu_j)}\left(1 - \sum_1^{k-1} y_j\right)^{\frac{1}{2}\nu_k - 1}\prod_{j=1}^{k-1} y_j^{\frac{1}{2}\nu_j - 1} \quad \left(0 \leqslant y_j;\ \sum_1^{k-1} y_j \leqslant 1\right), \quad (5a)'$$

which is of exactly the same form as $(5a)$, except that ν_j is replaced by $2(N-1)\pi_j$ for each $j = 1, 2, \ldots, k$. (It may be noted that Pearson (1923) states that L. N. G. Filon had, in 1901, suggested $(5a)$ as a possible 'multivariate frequency surface' for the joint distribution of a number of Pearson Type I variables. The same distribution was obtained by van Uven (1948) among his bivariate extensions of the Pearson system of curves.)

Hence we may regard our approximation to the joint distribution of $f_1, f_2, \ldots, f_k$ as equivalent to the replacement of these variables by

$$y_1 = v_1 \Big/ \sum_1^k v_j, y_2 = v_2 \Big/ \sum_1^k v_j, \ldots, y_k = v_k \Big/ \sum_1^k v_j,$$

respectively, the v's being independent χ^2 variables, v_j having $2(N-1)\pi_j$ degrees of freedom.

Equation $(5e)$ then follows from the additive property of independent χ^2's since $\sum_1^s v_j$ is distributed as χ^2 with $2(N-1)\sum_1^s \pi_j$ degrees of freedom.

5. The analogy described in the last section is useful when forming approximations to the distributions of certain statistics constructed from multinomial variables.

If the classes of the multinomial distribution are assigned fixed numerical values $X_1, X_2, \ldots, X_k$ we have a model which can apply to a frequency distribution (possibly grouped). The distribution of the rth crude moment

$$m_r' = \sum_{j=1}^k f_j X_j^r$$

would then be approximated by that of

$$\left(\sum_{j=1}^k X_j^r v_j\right) \Big/ \left(\sum_{j=1}^k v_j\right),$$

i.e. the ratio of a weighted sum of independent χ^2's to the unweighted sum of the same χ^2's.

In the case when the degrees of freedom are integers (i.e. π_j is an integer multiple of $\frac{1}{2}(N-1)^{-1}$) the ratio is of the form

$$\left(\sum_{r=1}^{2(N-1)} \lambda_r u_r^2\right) \Big/ \left(\sum_{r=1}^{2(N-1)} u_r^2\right),$$

where $u_1, u_2, \ldots, u_{2(N-1)}$ are independent unit Normal variables and the λ's are constants. The distribution of random variables of this type can be investigated by the methods used by von Neumann (1941).

Other consequences of the analogy are

(*a*) the ratio of the frequency in the ith cell to the frequency in the jth cell is approximately distributed as $(\pi_i/\pi_j) \times (F$ with $2(N-1)\pi_i$, $2(N-1)\pi_j$ degrees of freedom).

(*b*) in the symmetrical case $\pi_1 = \pi_2 = \ldots = \pi_k = k^{-1}$ the distribution of the ratio $(\max f_j)/(\min f_j) = (\max n_j)/(\min n_j)$ is approximated by that of the ratio $s^2_{\max}/s^2_{\min}$, for which some percentage points are given in Table 31 of Pearson & Hartley (1958).

(*c*) In the symmetrical case of (*b*) the maximum relative frequency $\max f_j$ has a distribution approximated by that of Cochran's criterion $s^2_{\max}/\sum_1^k s_j^2$, for which some percentage points are given in Eisenhart, Hastay & Wallis (1947).

Some numerical investigations of (*b*) and (*c*) are given in Johnson & Young (1960).

(*d*) The analogy does not lead to helpful results in the case of the statistic

$$\chi^2 = \sum_{j=1}^{k} (N/\pi_j^0)(f_j - \pi_j^0)^2$$

which is used to test the hypothesis H_0: $\pi_j = \pi_j^0 \quad (j = 1, 2, \ldots, k)$. We have

$$\chi^2 \doteqdot \frac{\sum_1^k (N/\pi_j^0) v_j^2}{\left(\sum_1^k v_j\right)^2} - N,$$

which is not easily handled analytically.

The likelihood ratio test of H_0 with respect to the alternatives that $\pi_j \neq \pi_j^0$ for at least two values of j leads to the criterion

$$M = -2(N-1) \sum_{j=1}^{k} f_j \log (f_j/\pi_j^0)$$

$$= -2(N-1) \sum_{j=1}^{k} \pi_j^0 (f_j/\pi_j^0) \log (f_j/\pi_j^0).$$

Large values of M are regarded as significant of departure from the hypotheses H_0. The distribution of M is not easy to estimate, even using our approximation. In § 6, approximate values of

$$\mu'_{\omega,\delta}(f, \log f; \pi) = \mathscr{E}\left(\prod_{j=1}^{k} f_j^{\omega_j} (\log f_j)^{\delta_j}\right)$$

(using an obvious notation) are obtained. These could be used to find, approximately, the moments of M and so to approximate to its distribution and to the power function of the test based on M.

Here we will only mention an alternative approach, using our analogy, which appears to be not unreasonable, applying the Bartlett–Neyman–Pearson test for equality of

variances (Neyman & Pearson (1931), Bartlett (1937)). The analogous criterion, in terms of the f_j's, is

$$M' = -2(N-1)\sum_{j=1}^{k} \pi_j^0 \log(f_j/\pi_j^0).$$

Large values of M' are regarded as significant of departure from the hypothesis H_0.

From Bartlett (1937)

$$M'\left[1+\frac{1}{6(k-1)(N-1)}\left\{\sum_{j=1}^{k}\frac{1}{\pi_j^0}-1\right\}\right]$$

is approximately distributed as χ^2 with $(k-1)$ degrees of freedom if H_0 is true.

M' is strictly appropriate to testing the hypothesis H_0 against the alternatives (corresponding to unequal variances, but the *same* degrees of freedom) specified by

$$g(f_1, f_2, \ldots, f_k) = \frac{\Gamma(N-1)}{\prod_{j=1}^{k}\Gamma((N-1)\pi_j^0)}\prod_{j=1}^{k}\{c_j(c_j f_j)^{(N-1)\pi_j^0-1}\} \quad \left(0 \leqslant f_j;\ \sum_1^k c_j f_j = 1\right) \qquad (2)'$$

rather than $\pi_j \neq \pi_j^0$, as stated above (which corresponds to changes in the number of degrees of freedom, the variances remaining the same). Nevertheless, the test based on M' should be fairly powerful with respect to hypotheses of type $\pi_j \neq \pi_j^0$ also, since the main effect, in both kinds of alternative hypotheses, is to change the ratio of the expected values of the f_j's.

6. We will suppose the values $f_j = 0$ omitted from our distributions; then we can consider approximations to the expected value of

$$\prod_{j=1}^{k}[f_j^{b_j}(\log f_j)^{\delta_j}].$$

We first note that the joint characteristic function of the k random variables $\log f_j$ is approximately

$$\begin{aligned}\phi_t(\log f) &= \mathscr{E}\left\{\exp\left(i\sum_1^k t_j \log f_j\right)\right\}\\ &= \mathscr{E}\left(\prod_{j=1}^{k} f_j^{it_j}\right)\\ &\doteqdot \frac{\Gamma(N-1)}{\Gamma\left(N-1+i\sum_1^k t_j\right)}\prod_{j=1}^{k}\frac{\Gamma((N-1)\pi_j+it_j)}{\Gamma((N-1)\pi_j)} \quad \text{from (3).}\end{aligned}$$

Hence, taking logarithms and differentiating, we have, for the cumulants of $\log f_j$

$$\kappa_r(\log f_j) \doteqdot \Psi^{(r-1)}((N-1)\pi_j) - \Psi^{(r-1)}(N-1), \qquad (7a)$$

where

$$\Psi^{(s)}(x) = \frac{d^{s+1}}{dx^{s+1}}\log\Gamma(x)$$

is the $(s+2)$-gamma function of x.

For the cross-cumulants of $\log f_{j_1}$ and $\log f_{j_2}$ $(j_1 \neq j_2)$ we obtain

$$\kappa_{r_1,r_2}(\log f_{j_1}, \log f_{j_2}) \doteqdot -\Psi^{(r_1+r_2-1)}(N-1), \qquad (7b)$$

and generally

$$\kappa_{r_1,r_2,\ldots,r_k}(\log f_1, \log f_2, \ldots, \log f_k) \doteqdot -\Psi^{(R-1)}(N-1), \qquad (7c)$$

where

$$R = \sum_{j=1}^{k} r_j.$$

Using equations (7) an approximation to any product moment

$$\mathscr{E}\left(\prod_{j=1}^{k}(\log f_j)^{\delta_j}\right) = \mu'_{\delta}(\log f;\pi) = \mu'_{0,\delta}(f,\log f;\pi)$$

(extending the notation introduced in § 6) can be found. In a number of cases explicit formulae have been obtained but they are not given here.

7. We have already seen (in § 2) that approximation (2) does not give correct multinomial moments of order higher than two. In this section we conclude our paper by noting two further types of difference from the exact model.

(*a*) Maximum likelihood estimators

The maximum likelihood estimator of π_j is f_j. If we use (2), which implies (5*d*), the maximum likelihood estimator $\hat{\pi}_j$, of π_j, satisfies the equation (in which $\Psi(x) = \Psi^{(0)}(x)$)

$$-(N-1)[\Psi((N-1)\hat{\pi}_j)-\Psi((N-1)(1-\hat{\pi}_j))]+(N-1)\log[f_j/(1-f_j)] = 0,$$

i.e.
$$\log\frac{f_j}{1-f_j} = \Psi((N-1)\hat{\pi}_j)-\Psi((N-1)(1-\hat{\pi}_j))$$
$$\doteqdot \log\left[\frac{(N-1)\hat{\pi}_j-\frac{1}{2}}{(N-1)(1-\hat{\pi}_j)-\frac{1}{2}}\right].$$

Hence
$$f_j[1-\hat{\pi}_j-\tfrac{1}{2}(N-1)^{-1}] \doteqdot (1-f_j)[\hat{\pi}_j-\tfrac{1}{2}(N-1)^{-1}]$$

and
$$\hat{\pi}_j \doteqdot f_j+\frac{1-2f_j}{2(N-1)},$$

$\hat{\pi}_j - f_j$ varies from $1/(2N)$ (approx.), when $f_j = 0$ to $-1/(2N)$ (approx.) when $f_j = 1$.

(*b*) Conditional distributions

The self-consistency of approximation (2) as exhibited in equations (5) does not extend to the conditional distributions of some f's, given others. Dividing (5*a*) by (5*b*) we get

$$p(f_{k-1}\,|\,f_1,f_2,\ldots,f_{k-2}) = \frac{\Gamma((N-1)(\pi_{k-1}+\pi_k))}{\Gamma((N-1)\pi_{k-1})\,\Gamma((N-1)\pi_k)}\left(\frac{f_{k-1}}{1-\sum_1^{k-2}f_j}\right)^{(N-1)\pi_{k-1}-1}$$
$$\times\left(1-\frac{f_{k-1}}{1-\sum_1^{k-2}f_j}\right)^{(N-1)\pi_k-1}\frac{1}{1-\sum_1^{k-2}f_j},$$

whence, making the transformation

$$f'_{k-1} = f_{k-1}\Big/\left(1-\sum_1^{k-2}f_j\right),$$

$$p(f'_{k-1}\,|\,f_1,f_2,\ldots,f_{k-2}) = \frac{\Gamma((N-1)(\pi_{k-1}+\pi_k))}{\Gamma((N-1)\pi_{k-1})\,\Gamma((N-1)\pi_k)}f'^{(N-1)\pi_{k-1}-1}_{k-1}(1-f'_{k-1})^{(N-1)\pi_k-1}$$
$$(0 \leqslant f'_{k-1} \leqslant 1). \quad (8a)$$

This means that, the conditional distribution of f'_{k-1} is of form (5*d*) (with $j = k-1$) with π_{k-1} replaced by $\pi_{k-1}/(\pi_{k-1}+\pi_k)$ and $(N-1)$ replaced by

$$(N-1)(\pi_{k-1}+\pi_k) = (N-1)\left(1-\sum_1^{k-2}\pi_j\right)$$

(i.e. N replaced by $N\left(1-\sum_1^{k-2}\pi_j\right)+\sum_1^{k-2}\pi_j$). For self-consistency N should have been replaced by

$$N\left(1-\sum_1^{k-2}f_j\right).$$

(Since $$\operatorname{var}\left(\sum_1^{k-2}f_j\right)=(\pi_{k-1}+\pi_k)(1-\pi_{k-1}-\pi_k)/N$$

the difference between the two new values of 'N' will 'usually' be of order $\sqrt{N}$.)

Similarly, in general

$$p(f_{a_1},f_{a_2},\ldots,f_{a_s}\,|\,f_{b_1},f_{b_2},\ldots,f_{b_t})=\frac{\Gamma\left((N-1)\left(1-\sum_1^t\pi_{b_l}\right)\right)}{\Gamma\left((N-1)\left(1-\sum_1^s\pi_{a_l}-\sum_1^t\pi_{b_l}\right)\right)\prod_1^s\Gamma((N-1)\pi_{a_l})}$$
$$\times\left(1-\frac{\sum_1^s f_{a_l}}{1-\sum_1^t f_{b_l}}\right)^{(N-1)(1-\Sigma\pi_{a_l}-\Sigma\pi_{b_l})-1}\prod_1^s\left(\frac{f_{a_l}}{1-\sum_1^t f_{b_l}}\right)^{(N-1)\pi_{a_l}-1}$$
$$\times\frac{1}{\left(1-\sum_1^t f_{b_l}\right)^s}\qquad\left(0\leqslant f_{a_l};\ \sum_1^s f_{a_l}\leqslant 1-\sum_1^t f_{b_l}\right).$$

Making the transformations

$$f''_{a_i}=f_{a_i}\Big/\left(1-\sum_1^t f_{b_l}\right)\quad(i=1,2,\ldots,s)$$

we find

$$p(f''_{a_1},f''_{a_2},\ldots,f''_{a_s}\,|\,f_{b_1},f_{b_2},\ldots,f_{b_t})=\frac{\Gamma\left((N-1)\left(1-\sum_1^t\pi_{b_l}\right)\right)}{\Gamma\left((N-1)\left(1-\sum_1^s\pi_{a_l}-\sum_1^t\pi_{b_l}\right)\right)\prod_1^s\Gamma((N-1)\pi_{a_l})}$$
$$\times\left(1-\sum_1^s f''_{a_l}\right)^{(N-1)(1-\Sigma\pi_{a_l}-\Sigma\pi_{b_l})-1}\prod_1^s f''^{(N-1)\pi_{a_l}-1}_{a_l}$$
$$\left(0\leqslant f''_{a_l};\ \sum_1^s f''_{a_l}\leqslant 1\right).\qquad(8b)$$

Here again, the relative probabilities

$$\pi_{a_l}\Big/\left(1-\sum_1^t\pi_{b_l}\right)$$

are correct, but (8b) implies a nominal 'sample size' equal to

$$N\left(1-\sum_1^t\pi_{b_l}\right)+\sum_1^t\pi_{b_l}$$

instead of the value

$$N\left(1-\sum_1^t f_{b_l}\right)$$

required for self-consistency.

REFERENCES

BARTLETT, M. S. (1937). Properties of sufficiency and statistical tests. *Proc. Roy. Soc.* A, **160**, 268–82.

BENEDETTI, C. (1956). Sulla rappresentabilità di una distribuzione binomiale mediante una distribuzione B e vice-versa. *Metron.* **18**, Nos. 1–2, 121–31.

DAVID, F. N. & JOHNSON, N. L. (1956). Reciprocal Bernoulli and Poisson variables. *Metron*, **18**, Nos. 1–2, 77–81.

DAVIES, O. L. (1933). On asymptotic formulae for the hypergeometric series. I. *Biometrika*, **25**, 295–322.

EISENHART, C., HASTAY, M. W. & WALLIS, W. A. (1947). *Selected Techniques of Statistical Analysis.* New York and London: McGraw-Hill.

FIELLER, E. C. & HARTLEY, H. O. (1954). Sampling with control variables. *Biometrika*, **41**, 494–501.

JOHNSON, N. L. & YOUNG, D. H. (1960). Some applications of two approximations to the multinomial distribution. (Submitted for publication in *Biometrika*.)

NEUMANN, J. VON (1941). Distribution of the ratio of the mean square successive difference to the variance. *Ann. Math. Statist.* **12**, 367–95.

NEYMAN, J. & PEARSON, E. S. (1931). On the problem of k samples. *Bull. Acad. Polon. Sci. Lett.* A, 460–81.

PEARSON, E. S. & HARTLEY, H. O. (1958). *Biometrika Tables for Statisticians*, **1**. Cambridge University Press.

PEARSON, K. (1923). Notes on skew frequency surfaces. *Biometrika*, **15**, 222–30.

UVEN, M. J. VAN (1948). Extension of Pearson's distributions to two variables. IV. *Proc. Kon. Ned. Akad. Wetensch.* (A), **51**, 191–6.

Biometrika (1960), **47**, 1 *and* 2, *p.* 103
Printed in Great Britain

The most economical binomial sequential probability ratio test

BY M. K. VAGHOLKAR AND G. B. WETHERILL
Bombay Textile Research Association and Birkbeck College, London

1. INTRODUCTION

Batches of items are presented for acceptance inspection, in which the items must be classified as either effective or defective. Let p denote the fraction defective contained in a batch. We assume that a critical fraction defective p_0 exists and is known, such that a batch with a fraction defective p_0 may without loss be accepted or rejected. We call batches with fraction defective $p < p_0$ good batches and these should be accepted. Those with fraction defective $p > p_0$ are termed bad batches and should be rejected. Let $P(p)$ denote the prior distribution (also called the process curve) for the fraction defective contained in batches submitted for inspection, and suppose this is of the following type

$$\left.\begin{aligned} &\Pr\{p = p_1\} = a_1, \quad \text{where} \quad a_1 + a_2 = 1, \\ &\Pr\{p = p_2\} = a_2 \quad \text{and} \quad p_1 < p_0 < p_2. \end{aligned}\right\} \tag{1}$$

Let W_{21} denote the decision cost incurred if a batch with fraction defective p_1 is rejected and let W_{12} denote the decision cost incurred if a batch with fraction defective p_2 is accepted. Let c denote the cost of inspecting a single item. The problem is to determine the most economical sampling inspection plan.

2. THE SEQUENTIAL PROBABILITY RATIO TEST

By assuming a special form of prior distribution, we have reduced our problem of acceptance sampling to that of testing two simple statistical hypotheses $\mathscr{H}_i$ $(i = 1, 2)$, where $\mathscr{H}_i$ means that a batch comes from a population with a fraction defective p_i $(i = 1, 2)$. We wish to accept a batch if $\mathscr{H}_1$ holds good, and if $\mathscr{H}_2$ holds, we wish to reject the batch.

In a sequential sampling inspection plan one item is inspected at a time, and the inspection is stopped as soon as sufficient evidence is obtained in favour of either of the hypotheses. Therefore as long as the cost of inspection depends merely on the total number of items inspected and no extra overhead cost is involved on account of sequential sampling, a sequential sampling inspection plan will be the most economical one. Moreover, Wald & Wolfowitz (1948) have shown that, when testing two simple hypotheses, of all tests with the same power, the sequential probability ratio test requires on average the fewest observations. Here this means that of all the sampling inspection plans with the same or lower decision costs, a plan based on the sequential probability ratio test will have the minimum cost of inspection. The usual sequential probability ratio test will therefore be the optimum test procedure, and this holds as long as the cost of inspection is linearly related to the number of items inspected. Below we give the method of arriving at the optimum test procedure, based on the basic theory given by Barnard (1954).

3. EQUATIONS FOR THE OPTIMUM TEST PROCEDURE

The optimum test procedure will be of the following type:

(*a*) continue inspection as long as

$$\lambda_2 < \lambda = \frac{a_1}{a_2} l(x, y) < \lambda_1; \tag{2}$$

(*b*) stop inspection and accept the batch as soon as $\lambda > \lambda_1$;

(*c*) stop inspection and reject the batch as soon as $\lambda < \lambda_2$;

where x and y denote respectively the number of effectives and defectives obtained at any stage, and $l(x, y)$ is the likelihood ratio and is equal to $(p_1/p_2)^y (q_1/q_2)^x$. It is more convenient to write inequality (2) as

$$L < l(x, y) < U, \tag{3}$$

where $U = a_2\lambda_1/a_1$ and $L = a_2\lambda_2/a_1$, and similar rules for acceptance and rejection apply.

Let $A(U, L; p)$ denote the probability that a batch with fraction defective p will be accepted;

$R(U, L; p)$ denote the probability that a batch with fraction defective p will be rejected;

and $S(U, L; p)$ denote the average number of items required to be inspected.

In order to obtain equations for the optimum boundaries we proceed in the manner indicated by Barnard (1954). The decision boundary is the locus of points such that the expected cost of taking an immediate decision is equal to the expected cost of taking at least one more observation and continuing the test. For example, when $\lambda = \lambda_1$, we write down the expected cost of accepting, as a function of the prior probabilities and decision costs: if we take another observation, then an effective item leads to accepting the batch, incurring a decision cost if $\mathscr{H}_2$ is true, and a defective item will take us to a point where we either continue the test or reject immediately.

We determine λ_1 to be that value for which these two costs—of accepting immediately, and of taking one more observation and continuing the test if necessary—are equal.

It should be noted that in general the difference between the expected cost of an immediate decision and the expected cost of continuing the test should decrease as the sample size increases.

For any λ, the posterior probabilities associated with the hypotheses $\mathscr{H}_1$ and $\mathscr{H}_2$ are $\lambda/(1+\lambda)$ and $1/(1+\lambda)$, respectively. We are now in a position to derive the equations for the optimum values λ_1 and λ_2.

Consider $\lambda = \lambda_1$, so that the probability that $\mathscr{H}_1$ is true is $\lambda_1(1+\lambda_1)$ and the probability that $\mathscr{H}_2$ is true is $1/(1+\lambda_1)$. Now the immediate decision to be taken is to accept $\mathscr{H}_1$, and we shall incur a loss W_{12} if $\mathscr{H}_2$ is true. The loss due to an immediate decision is therefore $W_{12}/(1+\lambda_1)$. Further if we inspect one more item, incurring a cost c, and continue the procedure thereafter, we shall accept the batch if the next item is an effective since $\lambda_1 q_1/q_2 > \lambda_1$. The cost W_{12} will be incurred if $\mathscr{H}_2$ is true, so that the loss is

$$W_{12} \times \{\text{prob. } \mathscr{H}_2 \text{ true}\} \times \{\text{prob. next item effective} \mid \mathscr{H}_2 \text{ true}\} = W_{12} q_2/(1+\lambda_1).$$

If the next item is a defective, the likelihood ratio will be $\lambda_1 p_1/p_2$, and we shall continue with a (new) sequential likelihood ratio test starting at this point. If the number of effectives and defectives obtained with this new test is (x', y'), the test is

$$\lambda_2 < \lambda_1 p_1 l(x', y')/p_2 < \lambda_1,$$

or

$$\frac{\lambda_2}{\lambda_1}\frac{p_2}{p_1} < l(x', y') < \frac{p_2}{p_1}.$$

Clearly if $\lambda_2/\lambda_1 \geqslant p_1/p_2$ we shall reject immediately, so that sampling continues only if $\lambda_2/\lambda_1 < p_1/p_2$. The cost due to this continuation is

$$[cS(\alpha, \rho\alpha; p_1) + W_{21}R(\alpha, \rho\alpha; p_1)]\, p_1\lambda_1/(1+\lambda_1)$$

if $\mathscr{H}_1$ is true and $\rho < p_1/p_2$, where $\alpha = p_2/p_1$, and $\rho = \lambda_2/\lambda_1$. If $\mathscr{H}_2$ is true and $\rho < p_1/p_2$ the continuation cost is

$$[cS(\alpha, \rho\alpha; p_2) + W_{12}A(\alpha, \rho\alpha; p_2)]p_2/(1+\lambda_1).$$

If $\rho > p_1/p_2$, we reject the batch, and the cost is $W_{21} \,.\, p_1 \,.\, \lambda_1/(1+\lambda_1)$.

We now equate the expected cost of an immediate decision to the expected cost if at least one more item is inspected, and we have,

$$W_{12}/(1+\lambda_1) = c + [cS(\alpha, \rho\alpha; p_1) + W_{21}R(\alpha, \rho\alpha; p_1)]\, p_1\lambda_1/(1+\lambda_1) + W_{12}q_2/(1+\lambda_1) + [cS(\alpha, \rho\alpha; p_2) + W_{12}A(\alpha, \rho\alpha; p_2)]\, p_2/(1+\lambda_1), \quad (4)$$

when $\rho < p_1/p_2$, and for $\rho > p_1/p_2$

$$W_{12}/(1+\lambda_1) = c + W_{21}p_1\lambda_1/(1+\lambda_1) + W_{12}q_2/(1+\lambda_1). \quad (5)$$

If we use the fact that $A(U, L; p) + R(U, L; p) = 1$ and simplify, equations (4) and (5) reduce to

$$\lambda_1 = \frac{W_{12}p_2R(\alpha, \rho\alpha; p_2) - c - p_2cS(\alpha, \rho\alpha; p_2)}{W_{21}p_1R(\alpha, \rho\alpha; p_1) + c + p_1cS(\alpha, \rho\alpha; p_1)} \quad \text{when} \quad \rho < p_1/p_2 \quad (6)$$

and

$$\lambda_1 = \frac{W_{12}p_2 - c}{W_{21}p_1 + c} \quad \text{when} \quad \rho > p_1/p_2. \quad (7)$$

A similar argument leads to the following equation for λ_2,

$$\lambda_2 = \frac{W_{12}q_2A(\beta|\rho, \beta; p_2) + c + q_2cS(\beta|\rho, \beta, p_2)}{W_{21}q_1A(\beta|\rho, \beta; p_1) - c - q_1cS(\beta|\rho, \beta; p_1)}, \quad \text{where} \quad \beta = q_2/q_1 \text{ and } \rho < q_2/q_1. \quad (8)$$

Dividing (8) by (6) we have for $\rho < p_1/p_2$

$$\rho = F(\rho) = \frac{[W_{12}q_2A(\beta|\rho, \beta; p_2) + c + q_2cS(\beta|\rho, \beta; p_2)]}{[W_{21}q_1A(\beta|\rho, \beta; p_1) - c - q_1cS(\beta|\rho, \beta; p_1)]} \times \frac{[W_{21}p_1R(\alpha, \rho\alpha; p_1) + c + p_1cS(\alpha, \rho\alpha; p_1)]}{[W_{12}p_2R(\alpha, \rho\alpha; p_2) - c - p_2cS(\alpha, \rho\alpha; p_2)]}. \quad (9)$$

Dividing (8) by (7) we have for $\rho > p_1/p_2$

$$\rho = F_1(\rho) = \frac{[W_{12}q_2A(\beta|\rho, \beta; p_2) + c + q_2cS(\beta|\rho, \beta; p_2)]\,[W_{21}p_1 + c]}{[W_{21}q_1A(\beta|\rho, \beta; p_1) - c - q_1cS(\beta|\rho, \beta; p_1)]\,[W_{12}p_2 - c]}. \quad (10)$$

We propose to solve equation (9) (or (10)) for ρ by iteration. The usual way of determining the boundaries of a sequential probability ratio test is to use Wald's approximate formulae which necessarily assume the risk of errors to be very small. This is not always true in the case of sampling inspection plans designed on a minimum cost basis. We therefore have to use some sort of exact formulae. The only such formulae that so far exist in the literature and which are practically manageable are those given by Burman (1946).

4. BARNARD'S SCORE NOTATION AND BURMAN'S SEQUENTIAL SAMPLING FORMULAE FOR A BINOMIAL POPULATION

According to Barnard (1946) any sequential probability ratio test procedure for a binomial population as defined by (3) can be reduced (without much loss in accuracy) to a scoring procedure as follows.

Take logarithms on both sides of (3) and divide throughout by $\log(q_1/q_2)$,

$$\frac{\log L}{\log(q_1/q_2)} < X - Y\frac{\log(p_2/p_1)}{\log(q_1/q_2)} < \frac{\log U}{\log(q_1/q_2)}.$$

Now put $$b = \frac{\log(p_2/p_1)}{\log(q_1/q_2)}, \quad H_1 = \frac{\log U}{\log(q_1/q_2)}, \quad H_2 = -\frac{\log L}{\log(q_1/q_2)}.$$

If we round off b, H_1 and H_2 to the nearest integers, the sequential likelihood ratio test procedure reduces to the following scoring scheme. Start with a score H_2, and add one to the score for each effective item found, subtracting b for each defective item found. Reject the batch if the score falls to zero or less, and accept the batch if the score reaches

$$2H = H_1 + H_2.$$

Formulae for the average sample size and acceptance probability of such a scheme have been given by Burman (1946), and they are exact if b, H_1 and H_2 are integers. The error involved in rounding off is small if b is greater than ten, and this is often the case in practice.

If we replace $A(U, L; p)$, etc., as introduced in § 3 by $A(H_1, H_2; p)$, etc., in score notation, equations (9) and (10) become

$$\rho = F(\rho) = \frac{[W_{12}q_2A(2H-1,1;p_2)+c+q_2cS(2H-1,1;p_2)]}{[W_{21}q_1A(2H-1,1;p_1)-c-q_1cS(2H-1,1;p_1)]} \times \frac{[W_{21}p_1R(b,2H-b;p_1)+c+p_1cS(b,2H-b;p_1)]}{[W_{12}p_2R(b,2H-b;p_2)-c-p_2cS(b,2H-b;p_2)]} \tag{11}$$

if $\rho < p_1/p_2$, and

$$\rho = F_1(\rho) = \frac{[W_{12}q_2A(2H-1,1;p_2)+c+q_2cS(2H-1,1;p_2)][W_{21}p_1+c]}{[W_{21}q_1A(2H-1,1;p_1)-c-q_1cS(2H-1,1;p_1)][W_{12}p_2-c]} \tag{12}$$

if $\rho > p_1/p_2$, where $$b = \frac{\log(p_2/p_1)}{\log(q_1/q_2)} \quad \text{and} \quad 2H = -\frac{\log\rho}{\log(q_1/q_2)}$$

rounded off to the nearest integer.

5. PROCEDURE FOR CALCULATION

In practice we shall first have to fix values for a_1, a_2, p_1, p_2, W_{12}, W_{21}, and c. The quantity b can then be calculated, or read directly from the *Target Handicap Charts* (1945). To solve equations (11) and (12) we start with some guessed value for ρ, and proceed with iteration until we get the same value of ρ. Actually on the right-hand side of equations (11) and (12), ρ enters through $2H$, which is an integer, so that when we get a ρ which gives rise to the same value of $2H$ as was used in the previous iteration, we stop and take that as our final solution. Once ρ is obtained we calculate λ_1 and λ_2 from equations (6) and (8) or (7) and (8) as the case may be.

The value of ρ is always less than one, and a lower limit for σ is derived by Vagholkar (1955) from which we have

$$c^2/[W_{12}(q_1-q_2)-c][W_{21}(q_1-q_2)-c] < \rho < 1. \tag{13}$$

From a number of examples that have been worked out so far it appears that the lower limit, or any number a little higher than the lower limit, can serve as a good first guess for ρ, in order to start the iteration.

6. Illustrative examples

Example 1. $a_1 = \frac{5}{9}$, $p_1 = 0{\cdot}01$, $W_{21} = 400$, $c = 1$;
$a_2 = \frac{4}{9}$, $p_2 = 0{\cdot}10$, $W_{12} = 500$.

We get $b = 24$, and by (13) we have $0{\cdot}00065 < \rho < 1$. Starting with a guess of $0{\cdot}001$ for ρ, the successive values of ρ obtained are,

$$\begin{array}{ccccc} \rho & \longrightarrow & 0{\cdot}001 & \longrightarrow & 0{\cdot}001607 \longrightarrow 0{\cdot}001600 \\ & & \downarrow & & \downarrow \quad\quad \downarrow \\ & & 2H = 72 & & 2H = 68 \quad\quad 2H = 68 \end{array}$$

This gives $\lambda_1 = 30{\cdot}5$, $\lambda_2 = 0{\cdot}0489$, and the optimum test procedure is given by

$$H_2 = 34, \quad 2H = H_1 + H_2 = 68, \quad b = 24.$$

Example 2. $a_1 = 0{\cdot}5981$, $p_1 = 0{\cdot}008663$, $W_{21} = 143{\cdot}6$, $c = 1$;
$a_2 = 0{\cdot}4019$, $p_2 = 0{\cdot}041856$, $W_{12} = 266{\cdot}0$.

Here $b = 46$, and the limits for ρ are

$$0{\cdot}0339 < \rho < 1.$$

Starting with a first guess of $\rho = 0{\cdot}1$ we finally get

$$\rho = 0{\cdot}2199, \quad \lambda_1 = 4{\cdot}5163 \quad \text{and} \quad \lambda_2 = 0{\cdot}9930.$$

In terms of score notation the optimum test procedure is

$$H_2 = 12, \quad 2H = H_1 + H_2 = 44 \quad \text{and} \quad b = 46.$$

7. The cost function

It has been assumed above that the cost of inspection is proportional to the number of items inspected, and that there are no other costs involved by the use of sequential sampling. Now in practice two particular types of departure from this assumption commonly arise, which can be taken account of as given below.

(i) In spite of the optimum properties of sequential schemes, they are not often used owing to the difficulty of inducing shop-floor staff to work them. The difficulty seems to arise in that a decision is required after every item inspected as to whether to continue sampling or sentence the batch. This can be taken account of in the cost function by adding an extra term which will depend on the number of times the decision to continue sampling or sentence the batch has to be made, which we denote D, and also on the cost associated with stopping sampling in order to make this decision, which we denote d. As a first

approximation we shall assume the extra term in the cost function to be the product of these two factors, Dd, though in practice a quadratic in D may be more realistic.

(ii) It often arises that items have to be processed before inspection. For example, in the climatic testing of resistors (see B.S. 2011, 1954) the items sampled are taken through temperature and humidity cycles, and the cost of processing a group of items is the same as the cost of processing one. We can take account of this by adding to the cost function the product of the number of times the processing is carried out, denoted T, with the cost associated with processing, denoted t, so that the product is denoted Tt.

By adding these two terms our cost function becomes

$$\text{cost} = nc + Dd + Tt. \tag{14}$$

The general problem for such cost functions will not be formulated here, but it is clear that the optimum test will be a likelihood ratio sequential test, with items sampled in groups (not necessarily of a constant size). For example, a certain valve factory sentences batches by inspecting an initial sample of 20 and succeeding samples (if any) of 5 valves. By limiting consideration to schemes with a constant group size, the problem is brought within the scope of the general theory, and we shall therefore do this. In defence it may be pointed out that practical considerations such as the capacity of the processing apparatus, or the need for simplicity may sometimes demand a constant group size. If the group size is k, equation (14) becomes

$$\begin{aligned} \text{cost} &= nc + \frac{n}{k}d + \frac{n}{k}t \\ &= nc', \quad \text{where} \quad c' = c + \frac{d}{k} + \frac{f}{k}, \end{aligned} \tag{15}$$

which is a linear function of the number of items inspected, with the artificial cost per item c'.

8. Calculation of group sampling boundaries

Two particular results apply to group sequential sampling, which immediately follow from the results of Wald & Wolfowitz (1948). They are stated here as lemmas.

Lemma 1. The risk of a group sequential sampling scheme is greater than or equal to the risk for a unit-step sequential, for given prior probability and loss functions.

Lemma 2. The optimum boundaries for group sequential sampling are within the optimum boundaries for unit-step sequential sampling with the same probability and loss functions.

The practical importance of these lemmas is that for group sequential sampling we can derive the optimum boundaries for unit-step sequential sampling using the artificial cost per item given in equation (15). These boundaries will contain the true optimum boundaries. A few trial calculations will result in approximately optimum boundaries. These trial calculations involve calculations of the group sampling continuation risk, which will have to be done numerically, truncating after two or three groups.

If the group size is large, and we can replace the two-point binomial distribution by an equivalent two-point normal distribution, a more exact solution is possible (Wetherill, 1959).

9. Multipoint prior distributions

So far this paper rests on approximating to a prior distribution by a two-point binomial. Vagholkar (1959) has suggested a method of doing this, and has given sampling schemes close to those given by the true prior distribution for examples so far calculated. Unfortunately, the case of more than two ordinates is too difficult to deal with mathematically at present, but we note the following result.

Consider, for example, the three-point binomial distribution given by

$$\Pr(p = p_i) = a_i \quad (i = 1, 2, 3), \tag{16}$$

where $\sum_{i=1}^{3} a_i = 1$ and $0 \leqslant p_1 < p_0 < p_2 \leqslant p_3 < 1$,

and p_0 is the critical fraction defective defined in § 1. The optimum decision boundaries for this prior distribution will in general meet (this holds for any number of ordinates greater than two). The equations are (Wetherill, 1957, p. 87)

$$a_1 p_1^r q_1^{n-r} W_{21} - a_2 p_2^r q_2^{n-r} W_{12} - a_3 p_3^r q_3^{n-r} W_{13} = 0, \tag{17}$$

$$a_1 p_1^r q_1^{n-r}(1 + p_1 W_{21}) + a_2 p_2^r q_2^{n-r}(1 - p_2 W_{12}) + a_3 p_3^r q_3^{n-r}(1 - p_3 W_{13}) = 0. \tag{18}$$

Equation (17) is the neutral line, or locus of points such that the risks of taking the two decisions are equal, and equation (18) is obtained by the condition that the two decision boundaries must have a common point. If $p_3 = p_2$, we have the two-ordinate case, and equation (18) reduces to

$$W_{12} W_{21}(p_2 - p_1) - W_{12} - W_{21} = 0, \tag{19}$$

which is the condition that it is not worth sampling (i.e. the two boundaries are identical) for the two-ordinate distribution based on (p_1, p_2).

It is found that equations (17) and (18) give a real solution if it is worth sampling for the two-ordinate (p_1, p_3) scheme, but that it is not worth sampling for the two-ordinate p_1, p_2 scheme.

The above work is based on theses submitted for the degree of Ph.D. in the University of London, and our thanks are due to Prof. G. A. Barnard of the Imperial College of Science, under whose guidance and supervision the work was done.

References

Barnard, G. A. (1946). Sequential tests in industrial statistics. *J. R. Statist. Soc.* Suppl. **8**, 1–27.

Barnard, G. A. (1954). Sampling inspection and statistical decisions. *J. R. Statist. Soc.* B, **16**, 151–74.

Burman, J. P. (1946). Sequential sampling formulae for a binomial population. *J. R. Statist. Soc.* Suppl. **8**, 98–103.

Target Handicap Charts for Sampling Inspection (1945). London: Ministry of Supply.

Vagholkar, M. K. (1955). Application of statistical decision theory to sampling inspection schemes. Ph.D. Thesis, University of London.

Vagholkar, M. K. (1959). The process curve and the equivalent mixed binomial with two components. *J. R. Statist. Soc.* B, **21**, 63–6.

Wald, A. & Wolfowitz, J. (1948). Optimum character of the sequential probability ratio test. *Ann. Math. Statist.* **19**, 326–39.

Wetherill, G. B. (1957). Application of the theory of decision functions to sampling inspection with special reference to cost of inspection. Ph.D. Thesis, University of London.

Wetherill, G. B. (1959). The most economical sequential sampling scheme for inspection by variables. *J. R. Statist. Soc.* B, **21**, 400–8.

Biometrika (1960), **47**, 1 *and* 2, *p.* 111
Printed in Great Britain

An approximate test for serial correlation in polynomial regression

BY J. R. McGREGOR
University of Birmingham

1. INTRODUCTION AND SUMMARY

The usual least squares methods of analysis of regression problems fail if the error terms in the regression model are serially correlated (see, for example, Durbin & Watson (1950) and Watson (1951)). Since the error terms themselves are not known, tests of serial correlation must be based on residuals from the fitted regression.

It has been customary to use test criteria of the form $\mathbf{z}'\mathbf{A}\mathbf{z}/\mathbf{z}'\mathbf{z}$, where $\mathbf{z}' = (z_1, z_2, \ldots, z_n)$ is the vector of residuals and $\mathbf{A}$ is a square matrix the choice of which depends on the type of correlation model to be tested. The only exact tests which have been developed are for special cases where the regression vectors coincide with certain latent vectors of the matrix $\mathbf{A}$. Thus R. L. & T. W. Anderson (1950) developed a method of finding the exact significance points of a circularly defined serial correlation coefficient based on the residuals from a fitted Fourier series.

Durbin & Watson (1950, 1951) developed bounds to the significance points of the test criterion

$$d = \frac{\sum_{j=1}^{n-1} (z_{j+1} - z_j)^2}{\sum_{j=1}^{n} z_j^2}. \tag{1·1}$$

The sample serial correlation may be defined in terms of d by the relation

$$r = 1 - \tfrac{1}{2}d. \tag{1·2}$$

The bounds to the significance points of d are 'best' in the sense that they are attained (when the regression vectors coincide with certain of the latent vectors of $\mathbf{A}$) and, when attained, the test criterion is uniformly most powerful against suitable alternative hypotheses. Hannan (1957) has remarked that, in the case of regression on the orthogonal polynomials, the upper bounds of Durbin & Watson differ from the true significance points by quantities of the order of n^{-2}, where n is the sample size.

In the present paper, an approach due to Daniels (1954, 1956) is used to obtain the approximate distribution of the sample serial correlation r (1·2). Hannan's remark on the proximity of the significance points of d (1·1) and d_u, the upper bound of Durbin & Watson's statistic, is verified and approximate significance points of d are compared with the corresponding significance points of d_u tabulated by Durbin & Watson (1951).

2. PRELIMINARY RESULTS

Before attacking the main problem of this paper, it will be convenient to review some preliminary results on the expansion of certain determinants.

Consider the $n \times n$ determinant

$$B_n = \begin{vmatrix} a & -1 & 0 & \dots & 0 & 0 \\ -1 & b & -1 & \dots & 0 & 0 \\ 0 & -1 & b & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & b & -1 \\ 0 & 0 & 0 & \dots & -1 & a \end{vmatrix}. \tag{2·1}$$

Let C_j be the determinant consisting of the first j rows and j columns of B_n ($j < n$). Then

$$C_1 = a, \quad C_2 = ab - 1 \quad \text{and} \quad C_j = bC_{j-1} - C_{j-2}, \quad 2 < j < n.$$

The solution of this difference equation is

$$C_j = \frac{(az-1)\,z^j}{z^2-1} + \frac{z(z-a)\,z^{-j}}{z^2-1},$$

where $z + 1/z = b$. Finally, since $B_n = aC_{n-1} - C_{n-2}$, we obtain

$$B_n = \frac{1}{z^n(1-z^2)}\{(z-a)^2 z^2 - z^{2n-2}(1-az)^2\}. \tag{2·2}$$

Now consider the $n \times n$ matrix

$$\mathbf{A} = \begin{bmatrix} 1-2T_0-T & -T & 0 & \dots & 0 & 0 \\ -T & 1-2T_0 & -T & \dots & 0 & 0 \\ 0 & -T & 1-2T_0 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 1-2T_0 & -T \\ 0 & 0 & 0 & \dots & -T & 1-2T_0-T \end{bmatrix}. \tag{2·3}$$

Factoring T from each row of $|\mathbf{A}|$ gives the form (2·1) whence, by (2·2),

$$|\mathbf{A}| = \frac{T^n}{z^n(1-z^2)}\left\{\left[z - \frac{(1-2T_0-T)}{T}\right]^2 z^2 - z^{2n-2}\left[1 - z\frac{(1-2T_0-T)}{T}\right]^2\right\}, \tag{2·4}$$

where

$$z + 1/z = (1-2T_0)/T. \tag{2·5}$$

The work of von Neumann (1941) leads to an alternative expansion of $|\mathbf{A}|$ in terms of the latent roots of the $n \times n$ matrix

$$\mathbf{Q} = \begin{bmatrix} 1 & -1 & 0 & \dots & 0 & 0 \\ -1 & 2 & -1 & \dots & 0 & 0 \\ 0 & -1 & 2 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 2 & -1 \\ 0 & 0 & 0 & \dots & -1 & 1 \end{bmatrix}.$$

von Neumann found the latent roots of $\mathbf{Q}$ to be

$$\lambda_j = 4\sin^2\frac{(j-1)\pi}{2n} \quad (j = 1, 2, \ldots, n), \tag{2·6}$$

so that the first root $\lambda_1 = 0$. We may write $\mathbf{A}$ in the form

$$\mathbf{A} = (1-2T_0-2T)\mathbf{I}+T\mathbf{Q}, \tag{2·7}$$

where $\mathbf{I}$ is the $n \times n$ identity matrix. Let $\mathbf{H}$ be an orthogonal matrix which diagonalizes $\mathbf{Q}$. Then

$$\mathbf{H}'\mathbf{A}\mathbf{H} = (1-2T_0-2T)\mathbf{I}+T\mathbf{H}'\mathbf{Q}\mathbf{H},$$

and thus

$$|\mathbf{A}| = (1-2T_0-2T)\prod_{j=2}^{n}(1-2T_0-2T+T\lambda_j). \tag{2·8}$$

It will be convenient to consider a third expansion of $|\mathbf{A}|$. Let $\mathbf{\Xi}$ be the $n \times n$ orthogonal matrix with elements

$$\xi_{ij} = \left\{\frac{(n-1)!}{(n+i-1)!}\right\}^{\frac{1}{2}}\frac{(2i-1)^{\frac{1}{2}}}{(i-1)!}\Delta^{i-1}(j-1)^{(i-1)}(j-1-n)^{(i-1)}, \tag{2·9}$$

where Δ is the forward difference operator defined by

$$\Delta f(x) = f(x+1)-f(x) \quad \text{and} \quad x^{(r)} = x(x-1)(x-2)\ldots(x-r+1)$$

with $x^{(0)} = 1$. From (2·7), we may write

$$\mathbf{\Xi}\mathbf{A}\mathbf{\Xi}' = (1-2T_0-2T)\mathbf{I}+T\mathbf{\Xi}\mathbf{Q}\mathbf{\Xi}'. \tag{2·10}$$

It is easily verified that $\mathbf{Q} = \mathbf{\Lambda}'\mathbf{\Lambda}$, where $\mathbf{\Lambda}$ is the $n \times n$ matrix

$$\mathbf{\Lambda} = \begin{bmatrix} 0 & 0 & 0 & \ldots & 0 & 0 \\ -1 & 1 & 0 & \ldots & 0 & 0 \\ 0 & -1 & 1 & \ldots & 0 & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & 0 & \ldots & 1 & 0 \\ 0 & 0 & 0 & \ldots & -1 & 1 \end{bmatrix},$$

so that $\mathbf{\Xi}\mathbf{Q}\mathbf{\Xi}' = (\mathbf{\Lambda}\mathbf{\Xi}')'\,\mathbf{\Lambda}\mathbf{\Xi}'$, where $\mathbf{\Lambda}\mathbf{\Xi}'$ has elements given by

$$\partial_{ij} = \xi_{ji}-\xi_{j,i-1}, \quad \partial_{i1} = \partial_{1j} = 0 \quad (i = 2, 3, \ldots, n;\ j = 2, 3, \ldots, n). \tag{2·11}$$

Thus $\mathbf{\Xi}\mathbf{Q}\mathbf{\Xi}' = \mathbf{D}$, where $\mathbf{D}$ is the $n \times n$ matrix with elements

$$d_{ij} = \sum_{s=2}^{n}\partial_{si}\partial_{sj}, \quad d_{i1} = d_{1j} = 0 \quad (i, j = 2, 3, \ldots, n). \tag{2·12}$$

Finally, from (2·10), we obtain

$$|\mathbf{A}| = |(1-2T_0-2T)\mathbf{I}+T\mathbf{D}|.$$

We note that, since $\mathbf{D}$ is obtained from $\mathbf{Q}$ by an orthogonal transformation, the latent roots of $\mathbf{D}$ and $\mathbf{Q}$ are the same and, by the invariance of the trace under orthogonal transformations,

$$\sum_{i=2}^{n}d_{ii} = \sum_{i=2}^{n}\sum_{k=2}^{n}\partial_{ki}^2 = \lambda_2+\lambda_3+\ldots+\lambda_n. \tag{2·13}$$

3. THE APPROXIMATE DISTRIBUTION OF THE SERIAL CORRELATION COEFFICIENT

We shall consider the regression model

$$x_s = \xi_{1s}\beta_1 + \xi_{2s}\beta_2 + \ldots + \xi_{ms}\beta_m + \epsilon_s \quad (s = 1, 2, \ldots, n),$$

where m is $O(1)$ with respect to n and the ϵ's are independent $N(0, 1)$ variables. The residuals z_s are

$$z_s = x_s - \sum_{i=1}^{m} \xi_{is} \sum_{k=1}^{n} \xi_{ik} x_k$$

$$= \epsilon_s - \sum_{i=1}^{m} \xi_{is} \sum_{k=1}^{n} \xi_{ik} \epsilon_k \quad (s = 1, 2, \ldots, n).$$

Transform to new independent variables y_s by $\mathbf{y} = \boldsymbol{\Xi}\boldsymbol{\epsilon}$, where $\boldsymbol{\Xi}$ is the $n \times n$ orthogonal matrix with elements ξ_{ij} (2·9) and $\mathbf{y}' = (y_1, y_2, \ldots, y_n)$, $\boldsymbol{\epsilon}' = (\epsilon_1, \epsilon_2, \ldots, \epsilon_n)$. Then

$$z_s = \sum_{i=m+1}^{n} \xi_{is} y_i, \tag{3·1}$$

with

$$\sum_{s=1}^{n} z_s^2 = \sum_{i=m+1}^{n} y_i^2 \tag{3·2}$$

by orthogonality. From (3·1), we have

$$z_{s+1} - z_s = \sum_{i=m+1}^{n} y_i(\xi_{i,s+1} - \xi_{i,s})$$

$$= \sum_{i=m+1}^{n} y_i \partial_{s+1,i},$$

where ∂_{ij} is defined by (2·11). Thus

$$\sum_{s=1}^{n-1} (z_{s+1} - z_s)^2 = \sum_{i=m+1}^{n} \sum_{j=m+1}^{n} y_i y_j \sum_{s=1}^{n-1} \partial_{s+1,i} \partial_{s+1,j}$$

$$= \sum_{i=m+1}^{n} \sum_{j=m+1}^{n} y_i y_j d_{ij}, \tag{3·3}$$

where d_{ij} is given by (2·12).

Let $\mathbf{D}_{n-m}$ be the $(n-m) \times (n-m)$ matrix with elements

$$d_{ij} = \sum_{s=1}^{n-1} \partial_{s+1,i} \partial_{s+1,j} = \sum_{s=2}^{n} \partial_{si} \partial_{sj} \quad (i, j = m+1, \ldots, n).$$

Then $\mathbf{D}_{n-m}$ is a principal submatrix of $\mathbf{D}$ (2·12). There is an orthogonal transformation $\mathbf{H}_{n-m}$ which will diagonalize $\mathbf{D}_{n-m}$ leaving the sums of squares invariant. Then

$$\mathbf{H}'_{n-m}\mathbf{D}_{n-m}\mathbf{H}_{n-m} = \begin{bmatrix} l_1 & & & (0) \\ & l_2 & & \\ & & \ldots & \\ (0) & & & l_{n-m} \end{bmatrix},$$

where $l_1, l_2, \ldots, l_{n-m}$ are the latent roots of $\mathbf{D}_{n-m}$. Define new variables $\zeta_1, \zeta_2, \ldots, \zeta_{n-m}$ by

$$\boldsymbol{\zeta} = \mathbf{H}'_{n-m}\mathbf{y}_{n-m},$$

where $\boldsymbol{\zeta}' = (\zeta_1, \zeta_2, \ldots, \zeta_{n-m})$ and $\mathbf{y}'_{n-m} = (y_{m+1}, y_{m+2}, \ldots, y_n)$.

Then, from (3·2) and (3·3),

$$\sum_{s=1}^{n} z_s^2 = \sum_{t=1}^{n-m} \zeta_t^2, \quad \text{and} \quad \sum_{s=1}^{n-1} (z_{s+1}-z_s)^2 = \sum_{t=1}^{n-m} \zeta_t^2 l_t.$$

Define the sample serial correlation to be $r = c/c_0$, where

$$c = \sum_{s=1}^{n} z_s^2 - \frac{1}{2}\sum_{s=1}^{n-1} (z_{s+1}-z_s)^2 = \sum_{t=1}^{n-m} \zeta_t^2 (1-\tfrac{1}{2}l_t),$$

$$c_0 = \sum_{s=1}^{n} z_s^2 = \sum_{t=1}^{n-m} \zeta_t^2.$$

The joint moment-generating function of c_0 and c is

$$M(T_0, T) = E(e^{T_0 c_0 + Tc}) = \prod_{t=1}^{n-m} (1-2T_0-2T+Tl_t)^{-\frac{1}{2}}. \tag{3·4}$$

It is shown in the Appendix that each $(\lambda_{m+s}-l_s)$ is positive and *at most* $O(n^{-2})$ $(s = 1, 2, \ldots, n-m)$, and that $\sum_{s=1}^{n-m} (\lambda_{m+s}-l_s)$ is $O(n^{-2})$. Thus

$$\prod_{t=1}^{n-m} (1-2T_0-2T+Tl_t) \sim \prod_{t=m+1}^{n} (1-2T_0-2T+T\lambda_t)\{1+O(n^{-2})\}. \tag{3·5}$$

Using equation (2·8), we obtain

$$\prod_{t=m+1}^{n} (1-2T_0-2T+T\lambda_t) = \frac{|\mathbf{A}|}{(1-2T_0-2T)\prod_{t=2}^{m} (1-2T_0-2T+T\lambda_t)}.$$

Since m is $O(1)$ with respect to n

$$\lambda_t = 4\sin^2\frac{(t-1)\pi}{2n} \quad \text{is of order } n^{-2} \quad (t = 2, 3, \ldots, m).$$

Hence,
$$\prod_{t=m+1}^{n} (1-2T_0-2T+T\lambda_t) \sim \frac{|\mathbf{A}|}{(1-2T_0-2T)^m}\{1+O(n^{-2})\},$$

and, from (3·5),
$$\prod_{t=1}^{n-m} (1-2T_0-2T+Tl_t) \sim \frac{|\mathbf{A}|}{(1-2T_0-2T)^m}\{1+O(n^{-2})\}.$$

Finally, using (2·4) and (2·5), we obtain

$$\prod_{t=1}^{n-m} (1-2T_0-2T+Tl_t) \sim \frac{T^{n-m}}{z^{n-m}(1-z)^{2m}(1-z^2)} \times \left\{\left[z-\frac{(1-2T_0-T)}{T}\right]^2 z^2 - z^{2n-2}\left[1-\frac{z(1-2T_0-T)}{T}\right]^2\right\}\{1+O(n^{-2})\}.$$

Since $|z| < 1$ except in critical cases where $|z| = 1$, we incur an exponentially small error in omitting the second term in the brackets. Then, from (3·4), we have

$$M(u-rT,T) \sim \frac{(1-2rz+z^2)^{\frac{1}{2}(n-m)}(1-z)^{m-1}(1-z^2)^{\frac{1}{2}}}{(1-2u)^{\frac{1}{2}(n-m)}}\{1+O(n^{-2})\},$$

and

$$\frac{\partial}{\partial u}\left[M(u-rT,T)\left(\frac{\partial T}{\partial z}\right)_u\right]\Big|_{u=0} \sim (n-m-2)(1-2rz+z^2)^{\frac{1}{2}(n-m)-2}(1-z)^{m-1}(1-z^2)^{\frac{1}{2}}\{1+O(n^{-2})\},$$

where we have made the substitutions

$$u = T_0+rT \quad \text{and} \quad z+1/z = (1-2T_0)/T.$$

To obtain the approximate distribution of r, we follow closely the arguments of Daniels (1956). The density function $h(r)$ of r is given approximately by

$$h(r) \sim \frac{(n-m-2)}{2\pi i}\int \phi(z)(1-2rz+z^2)^{\frac{1}{2}(n-m)-2}\,dz,$$

where

$$\phi(z) = (1-z)^{m-1}(1-z^2)^{\frac{1}{2}},$$

the path of integration being the same as that described by Daniels (1956, p. 172) for the case $\rho = 0$. Expanding $\phi(z)$ in Taylor's series about $z = r$ and integrating, we obtain the first approximation

$$h(r) \sim \frac{(n-m-2)}{2\pi^{\frac{1}{2}}}\frac{\Gamma(\frac{1}{2}(n-m)-1)}{\Gamma(\frac{1}{2}(n-m)-\frac{1}{2})}(1-r)^{m-1}(1-r^2)^{\frac{1}{2}(n-m)}\{1+O(n^{-1})\}. \tag{3·6}$$

In re-normalized form, (3·6) becomes

$$h(r) \sim \frac{(1-r)^{m-1}(1-r^2)^{\frac{1}{2}(n-m)}}{2^n B[\frac{1}{2}(n-m)+1,\frac{1}{2}(n+m)]}\{1+O(n^{-2})\}, \tag{3·7}$$

where $B(p,q)$ is the usual β-function and where an argument of Daniels (1956, p. 178) shows that the relative error is $O(n^{-2})$.

The serial correlation r is related to the statistic d (1·1) of Durbin & Watson (1950, 1951) by the relation $r = 1-\frac{1}{2}d$. Using this relation, the approximate renormalized density function of d is readily found to be

$$g(d) \sim \frac{d^{\frac{1}{2}(n+m)-1}(4-d)^{\frac{1}{2}(n-m)}}{2^{2n}B[\frac{1}{2}(n-m)+1,\frac{1}{2}(n+m)]}\{1+O(n^{-2})\} \quad (0 \leqslant d \leqslant 4). \tag{3·8}$$

It is shown in the Appendix that the exact distribution of d, in the case of regression on the orthogonal polynomials, differs from that of d_u, the upper bound of Durbin & Watson's statistic, by quantities which are relatively $O(n^{-2})$. It follows that the corresponding approximate distribution of d (3·8) differs from the distribution of d_u by quantities of the same order. Table 1 compares 5 % and 1 % significance points of d calculated from (3·8) with the corresponding points tabulated by Durbin & Watson (1951) for d_u, for certain values of m and n.

Table 1. *5 % and 1 % significance points of d (approximate) and d_u*

$k' = m-1$ →		1		3		5	
n		d	d_u	d	d_u	d	d_u
16	5 %	1·33	1·37	1·56	1·73	1·80	2·15
	1 %	1·04	1·09	1·25	1·44	1·48	1·90
20	5 %	1·39	1·41	1·58	1·68	1·77	1·99
	1 %	1·12	1·15	1·29	1·41	1·48	1·74
24	5 %	1·43	1·45	1·59	1·66	1·75	1·90
	1 %	1·18	1·20	1·33	1·41	1·48	1·66
30	5 %	1·48	1·49	1·61	1·65	1·73	1·83
	1 %	1·25	1·26	1·37	1·42	1·50	1·61
40	5 %	1·54	1·54	1·64	1·66	1·73	1·79
	1 %	1·34	1·34	1·43	1·46	1·52	1·58

I am indebted to Prof. H. E. Daniels for his advice during the preparation of this paper.

Appendix

Consider the diagonal elements of **D** defined by (2·12). These are

$$d_{11} = 0, \quad d_{ii} = \sum_{j=2}^{n} \partial_{ji}^2 \quad (i = 2, 3, \ldots, n),$$

where $\partial_{ji} = \xi_{ij} - \xi_{i,j-1} = \Delta\xi_{i,j-1}$ with Δ the forward difference operator on j and ξ_{ij} defined by (2·9). Then

$$d_{ii} = \sum_{j=2}^{n} (\Delta\xi_{i,j-1})^2$$

$$= \frac{(n-i)!}{(n+i-1)!} \frac{(2i-1)}{[(i-1)!]^2} \sum_{j=2}^{n} \{\Delta^i (j-2)^{(i-1)} (j-2-n)^{(i-1)}\}^2 \quad (i = 2, 3, \ldots, n). \tag{A 1}$$

Let $x = j-2$ and $r = i-1$. Then

$$\sum_{j=2}^{n} \{\Delta^i (j-2)^{(i-1)} (j-2-n)^{(i-1)}\}^2 = \sum_{x=0}^{n-2} \{\Delta^{r+1} x^{(r)} (x-n)^{(r)}\}^2. \tag{A 2}$$

Adopting the notation of Szegö (1939), we write this as

$$(r!)^2 \sum_{x=0}^{n-2} (\Delta t_r(x))^2 = \sum_{x=0}^{n-2} [\Delta^{r+1} x^{(r)} (x-n)^{(r)}]\, \Delta[\Delta^r x^{(r)} (x-n)^{(r)}], \tag{A 3}$$

where

$$t_r(x) = \frac{1}{r!} \Delta^r x^{(r)} (x-n)^{(r)}.$$

Summing by parts on the right of (A 3), we obtain

$$(r!)^2 \sum_{x=0}^{n-2} (\Delta t_r(x))^2 = [(\Delta^{r+1} x^{(r)} (x-n)^{(r)})\, \Delta(\Delta^r x^{(r)} (x-n)^{(r)})]_{x=0}^{x=n-1}$$

$$- \sum_{x=0}^{n-2} (\Delta^{r+2} x^{(r)} (x-n)^{(r)})\, \Delta(\Delta^{r-1} (x+1)^{(r)} (x+1-n)^{(r)}).$$

Performing repeated summations by parts on the right and noting that $\Delta^{2r+k} x^{(r)} (x-n)^{(r)} = 0$, $k > 0$, we ultimately obtain

$$(r!)^2 \sum_{x=0}^{n-2} (\Delta t_r(x))^2 = \sum_{s=0}^{r-1} (-1)^s [(\Delta^{r+s+1} x^{(r)} (x-n)^{(r)})\, (\Delta^{r-s} (x+s)^{(r)}\, (x+s-n)^{(r)})]_{x=0}^{x=n-1}. \tag{A 4}$$

It is easily verified that

$$\Delta^{r-s}(x+s)^{(r)}(x+s-n)^{(r)}\big|_{x=n-1} = 0 \quad (s>0),$$

$$= \frac{r!\,(n-1)!}{(n-1-r)!} \quad (s=0),$$

and that $$\Delta^{r-s}(x+s)^{(r)}(x+s-n)^{(r)}\big|_{x=0} = \frac{(-1)^r r!\,(n-1)!}{(n-1-r)!}.$$

Hence (A 4) may be written,

$$(r!)^2\sum_{x=0}^{n-2}(\Delta t_r(x))^2 = \frac{r!\,(n-1)!}{(n-1-r)!}\left\{[\Delta^{r+1}x^{(r)}(x-n)^{(r)}]_{x=n-1} - \sum_{s=0}^{r-1}(-1)^{s+r}[\Delta^{r+s+1}x^{(r)}(x-n)^{(r)}]_{x=0}\right\}. \quad (A\,5)$$

It may also be verified that

$$-\sum_{s=0}^{r-1}(-1)^{s+r}[\Delta^{r+s+1}x^{(r)}(x-n)^{(r)}]_{x=0} = (-1)^{r+1}[\Delta^{r+1}x^{(r)}(x-n)^{(r)}]_{x=-1}.$$

Thus $$(r!)^2\sum_{x=0}^{n-2}(\Delta t_r(x))^2 = \frac{r!\,(n-1)!}{(n-1-r)!}\{[\Delta^{r+1}x^{(r)}(x-n)^{(r)}]_{x=n-1} + (-1)^{r+1}[\Delta^{r+1}x^{(r)}(x-n)^{(r)}]_{x=-1}\}. \quad (A\,6)$$

Using Cauchy's integral formula, we may write,

$$\Delta^{r+1}x^{(r)}(x-n)^{(r)} = \frac{(r+1)!}{2\pi i}\int_C \frac{z(z-1)\dots(z-r+1)(z-n)(z-n-1)\dots(z-n-r+1)}{(z-x)(z-x-1)\dots(z-x-r-1)}\,dz, \quad (A\,7)$$

where the points $x, x+1, \dots, x+r+1$ are interior to the simple closed curve C. In particular, for $x = n-1$,

$$[\Delta^{r+1}x^{(r)}(x-n)^{(r)}]_{x=n-1} = \frac{(r+1)!}{2\pi i}\int_C \frac{z(z-1)\dots(z-r+1)}{(z-n+1)(z-n-r)}\,dz. \quad (A\,8)$$

We may evaluate (A 8) by residues and obtain

$$[\Delta^{r+1}x^{(r)}(x-n)^{(r)}]_{x=n-1} = r!\left[\frac{(n+r)!}{n!} - \frac{(n-1)!}{(n-1-r)!}\right]. \quad (A\,9)$$

Similarly, for $x = -1$, we get

$$[\Delta^{r+1}x^{(r)}(x-n)^{(r)}]_{x=-1} = (-1)^{r+1}r!\left[\frac{(n+r)!}{n!} - \frac{(n-1)!}{(n-1-r)!}\right]. \quad (A\,10)$$

Finally, from (A 6), (A 9) and (A 10),

$$(r!)^2\sum_{x=0}^{n-2}(\Delta t_r(x))^2 = \frac{2(r!)^2(n-1)!}{(n-1-r)!}\left[\frac{(n+r)!}{n!} - \frac{(n-1)!}{(n-1-r)!}\right],$$

$$= \frac{2[(i-1)!]^2(n-1)!}{(n-i)!}\left[\frac{(n+i-1)!}{n!} - \frac{(n-1)!}{(n-i)!}\right], \quad (A\,11)$$

where $r = i-1$. From (A 1), (A 2) and (A 11), we get

$$d_{ii} = \frac{2(2i-1)(n-1)!}{(n+i-1)!}\left[\frac{(n+i-1)!}{n!} - \frac{(n-1)!}{(n-i)!}\right] \quad (i = 2, 3, \dots, n). \quad (A\,12)$$

Now consider the $(n-m)\times(n-m)$ matrix $\mathbf{D}_{n-m}$ with elements d_{ij} $(i = m+1, \dots, n;\ j = m+1, \dots, n)$. $\mathbf{D}_{n-m}$ is the principal submatrix of $\mathbf{D}$ located in the lower right corner. The latent roots of $\mathbf{D}$ are $\lambda_1 = 0 < \lambda_2 < \dots < \lambda_n$, given by (2·6). There will be no loss of generality in taking the latent roots of $\mathbf{D}_{n-m}$ to be $l_1 \leqslant l_2 \leqslant \dots \leqslant l_{n-m}$. By (2·13), we have

$$\lambda_2 + \lambda_3 + \dots + \lambda_n = (d_{22} + \dots + d_{mm}) + (d_{m+1,m+1} + \dots + d_{nn}). \quad (A\,13)$$

Since the trace is invariant under orthogonal transformations, $d_{m+1,m+1} + \dots + d_{nn} = l_1 + l_2 + \dots + l_{n-m}$. Thus from (A 13) we obtain

$$\lambda_2 + \lambda_3 + \dots + \lambda_n = (d_{22} + \dots + d_{mm}) + l_1 + l_2 + \dots + l_{n-m}, \quad (A\,14)$$

and hence,

$$(\lambda_n - l_{n-m}) + (\lambda_{n-1} - l_{n-m-1}) + \dots + (\lambda_{m+1} - l_1) = (d_{22} + \dots + d_{mm}) - (\lambda_2 + \dots + \lambda_m). \quad (A\,15)$$

It follows from the spacing of the latent roots of principal submatrices of real symmetric matrices (see, for example, Turnbull & Aitken (1932)), that each term on the left of (A 15) is positive or zero. Now consider the terms on the right of (A 15). Since m is $O(1)$ with respect to n,

$$0 < \lambda_2 < \lambda_3 < \ldots < \lambda_m = 4\sin^2\frac{(m-1)\pi}{2n} = O(n^{-2}).$$

Thus $\lambda_2 + \ldots + \lambda_m$ is $O(n^{-2})$. From (A 12),

$$d_{ii} = \frac{2(2i-1)\,[(n+i-1)\,(n+i-2)\ldots(n+1)-(n-1)\,(n-2)\ldots(n-i+1)]}{(n+i-1)\,(n+i-2)\ldots(n+1)\,n} \quad (i = 2, 3, \ldots, n).$$

In particular, for $2 \leqslant i \leqslant m$, d_{ii} is $O(n^{-2})$. Thus $(\lambda_n - l_{n-m}) + (\lambda_{n-1} - l_{n-m-1}) + \ldots + (\lambda_{m+1} - l_1)$ is $O(n^{-2})$. Since each term in this sum is positive, it follows that each $(\lambda_{m+s} - l_s)$ is *at most* $O(n^{-2})$ $(s = 1, 2, \ldots, n-m)$. Thus to $O(n^{-2})$ we may replace l_s by λ_{m+s} $(s = 1, 2, \ldots, n-m)$.

Durbin & Watson's statistic d_u may be written $d_u = \sum_{i=1}^{n-m} \lambda_{m+i}\zeta_i^2 \Big/ \sum_{i=1}^{n-m} \zeta_i^2$, where the ζ's are independent $N(0, 1)$ variables. The joint moment-generating function of the numerator and denominator of d_u is

$$M_1(T_0, T) = E\left\{\exp\left[T_0 \sum_{i=1}^{n-m} \zeta_i^2 + T \sum_{i=1}^{n-m} \lambda_{m+i}\zeta_i^2\right]\right\}$$

$$= \prod_{i=1}^{n-m} [1 - 2T_0 - 2T\lambda_{m+i}]^{-\frac{1}{2}}.$$

The significance points of the distribution of d_u give upper bounds to the significance points of the statistic d. For regression on the orthogonal polynomials, d may be written as

$$d = \sum_{i=1}^{n-m} l_i\zeta_i^2 \Big/ \sum_{i=1}^{n-m} \zeta_i^2,$$

and hence the joint moment-generating function of the numerator and denominator of d is

$$M_2(T_0, T) = \prod_{i=1}^{n-m} (1 - 2T_0 - 2Tl_i)^{-\frac{1}{2}}.$$

It follows from the foregoing remarks that, in the case of the orthogonal polynomials, the distribution of d differs from that of d_u by quantities which are $O(n^{-2})$ since, to this order, we may replace l_i with λ_{m+i}. This conclusion was reached independently by Hannan (1957).

References

Anderson, R. L. & Anderson, T. W. (1950). Distribution of the circular serial correlation coefficient for residuals from a fitted Fourier series. *Ann. Math. Statist.* **21**, 59–81.

Daniels, H. E. (1954). Saddlepoint approximations in statistics. *Ann. Math. Statist.* **25**, 631–50.

Daniels, H. E. (1956). The approximate distribution of serial correlation coefficients. *Biometrika*, **43**, 169–85.

Durbin, J. & Watson, G. S. (1950). Testing for serial correlation in least squares regression. I. *Biometrika*, **37**, 409–28.

Durbin, J. & Watson, G. S. (1951). Testing for serial correlation in least squares regression. II. *Biometrika*, **38**, 159–78.

Hannan, E. J. (1957). Testing for serial correlation in least squares regression. *Biometrika*, **44**, 57–66.

Neumann, J. von (1941). Distribution of the ratio of the mean square successive difference to the variance. *Ann. Math. Statist.* **12**, 367–95.

Szegö, G. (1939). *Orthogonal Polynomials.* New York: Mathematical Society Colloquium Publications, **23**.

Turnbull, H. W. & Aitken, A. C. (1932). *An Introduction to the Theory of Canonical Matrices.* London and Glasgow: Blackie and Son Ltd.

Watson, G. S. (1951). *Serial Correlation in Regression Analysis.* Ph.D. Thesis (unpublished): University of North Carolina.

Biometrika (1960), **47**, 1 *and* 2, *p.* 121
Printed in Great Britain

Nomograms for fitting the logistic function by maximum likelihood

BY JOSEPH BERKSON, M.D., D.Sc.*
Section of Biometry and Medical Statistics, Mayo Clinic, Rochester, Minnesota

1. THEORY AND DESCRIPTION OF THE TABLES

The logistic function has been used as a model for bio-assay and other experiments in which the relative frequency, p, corresponding to the probability P of some 'response' is observed for some 'stimulus', x, and P is a monotonic function of x. The 'response' can be, for example, the deaths among n animals to which some mortally toxic drug of dosage x has been administered, or the explosions among n shells that have been subjected to a percussion of force x.

The logistic function can be written

$$P_i = 1 - Q_i = \frac{1}{1+e^{-(\alpha+\beta x_i)}}, \tag{1}$$

where P_i is the probability of the response at x_i. At each of $k \geqslant 2$ values of x, assumed to be without error, the relative frequency $p_i = r_i/n_i$, assumed to be a binomial variable on P_i, is observed, where r_i is the number of responses out of n_i trials. The parameters to be estimated are α, β. Frequently it is desired to estimate the ED 50, which is the value of x corresponding to 100 $P_i = 50$ %; this is given by $\gamma = -\alpha/\beta$.

The maximum likelihood estimates of α, β represented as $\hat{\alpha}, \hat{\beta}$, are given by the solution for $\hat{\alpha}, \hat{\beta}$, of the estimating equations

$$\sum_{i=1}^{k} n_i p_i = \sum_{i=1}^{k} n_i \hat{p}_i, \tag{2}$$

$$\sum_{i=1}^{k} n_i p_i x_i = \sum_{i=1}^{k} n_i \hat{p}_i x_i, \tag{3}$$

where $\hat{p}_i$, is the estimate of P_i given by

$$\hat{p}_i = \frac{1}{1+e^{-(\hat{\alpha}+\hat{\beta}x_i)}}. \tag{4}$$

The solutions of (2) and (3) cannot in general be obtained explicitly, and an iterative procedure must be used which yields $\hat{\alpha}, \hat{\beta}$, as limits, when these limits exist (Berkson, 1957). The maximum likelihood estimate of γ is $\hat{\gamma} = -\hat{\alpha}/\hat{\beta}$.

However, it can be seen that the estimates are functions of

$$\Sigma n_i p_i = \Sigma r_i \quad \text{and} \quad \Sigma n_i p_i x_i = \Sigma r_i x_i.$$

$\Sigma n_i p_i$ is minimal sufficient for α, with β fixed, and $\Sigma n_i p_i x_i$ is minimal sufficient for β, with α fixed (Berkson, 1955). The maximum likelihood estimates of α and β are functions of $\Sigma n_i p_i$, $\Sigma n_i p_i x_i$ which is illustrative of the theorem, proved early in the history of the subject, that when a minimal sufficient set of statistics other than the complete observations of the sample itself exists, the maximum likelihood estimate is a function of these

* This project was supported in part by the National Science Foundation.

sufficient statistics (Fisher, 1925; Neyman, 1935). If all $n_i = n$, then the maximum likelihood estimates of α, β, will be functions of Σp_i, $\Sigma p_i x_i$ independent of the value of n.

Worcester & Wilson (1943) noted in the case of three equally spaced doses with equal n, that the maximum likelihood estimates can be tabled against an argument of two quantities which, when examined, can be seen to be in one-to-one relation with Σp_i, $\Sigma p_i x_i$, and these authors provided such a table. Cornfield & Mantel (1950) pointed out that such tables could be constructed for any specified dosage arrangement, and the present author (Berkson, 1953) observed the same fact but suggested and illustrated nomographic presentation. For the required refinement of reading of argument, tabular presentation is formidably elaborate. Moreover, computation of such tables requires calculations which are quite tedious. For preparation of nomograms giving the estimate of the parameters from Σp, Σpx, we use the fact that for any value of α (or γ) with varying β, the corresponding values of Σp, Σpx are the same values which, if found from a sample, yield as the maximum likelihood estimate the α (or γ) in question. Similarly, one can determine the values of Σp, Σpx corresponding to a defined value of β with varying α.

The nomograms presented in the following pages provide maximum likelihood estimates of γ and β from computed values of Σp and Σpx, for equally spaced dosages and equal n at each dose, for 3, 4, 5 and 6 doses.* A separate nomogram is given for γ and for β.

For the most efficient spacing in the construction of the nomograms the dosages are coded with $x = 0$ at the centre of the range. But for computation, it is convenient to scale the dosages as $0, 1, 2, \ldots, (k-1)$. The dosage on the computing scale will be referred to as X, that on the nomographic scale as x. For the use of the nomograms, one computes Σp and ΣpX from the data recorded at dosages X. Then, for consulting the nomogram

$$\Sigma px = \Sigma pX - \tfrac{1}{2}(k-1)\Sigma p. \tag{5}$$

The value of γ is given in the nomogram directly in terms of the computing scale X. In order to increase the sensitivity of the nomograms they are designed to give γ only for positive values in terms of x. When the value of γ in these terms is positive Σp will be located on the left-hand scale, when it is negative Σp will be found on the right-hand scale. In the latter case the value read as γ in terms of X is converted to the correct value by

$$\hat{\gamma} = k - 1 - \gamma. \tag{6}$$

Values of $\hat{\gamma}$ are provided over the entire, or almost the entire, range of doses. In a well-designed experiment, the ED 50 should fall near the centre of the range of doses; if it falls far from this position the experiment should be redone. Values of β are given over a range found to cover practical work.

For the variances of the estimates, the following 'large sample' formulae are available:

$$s^2(\hat{\beta}) = \frac{1}{n\Sigma\hat{p}_i\hat{q}_i(x_i-\bar{x})^2}, \tag{7}$$

$$s^2(\hat{\gamma}) = \frac{1}{n\hat{\beta}^2}\left[\frac{1}{\Sigma\hat{p}_i\hat{q}_i} + \frac{(\hat{\gamma}-\bar{x})^2}{\Sigma\hat{p}_i\hat{q}_i(x_i-\bar{x})^2}\right], \tag{8}$$

where $\bar{x} = \Sigma n_i\hat{p}_i\hat{q}_i x_i / \Sigma n_i\hat{p}_i\hat{q}_i$.

* A nomogram was prepared also for 2 doses and plans for its publication announced (Hodges, 1958). However, since for this experiment the estimates of the parameters are given by the straight line join of the points represented on a logit graph, or the simple addition and subtraction of the corresponding logits, it was decided that its publication would be supererogatory.

Since, for any specified dosage arrangement, $\hat{p}$, $\hat{q}$ and $\bar{x}$ are determined by $\hat{\gamma}$, $\hat{\beta}$, the estimated variances are determined by the estimates $\hat{\gamma}$, $\hat{\beta}$ and n. In the tables preceding the nomograms the standard error of each estimate, multiplied by $\sqrt{n}$, is given to two significant figures, which are enough for practical purposes. The standard errors change slowly with $\hat{\gamma}$, $\hat{\beta}$, which are tabled as argument only with sufficient refinement to determine the standard errors unequivocally.

2. Illustrations

Example 1 ($k = 4$ *doses*).

X	n	r	p		
0	30	1	0·0333	$\Sigma p = 1{\cdot}5667,$	$\Sigma px = \Sigma pX - 1{\cdot}5\Sigma p = 1{\cdot}21675.$
1	30	8	0·2667	$\Sigma pX = 3{\cdot}5668,$	
2	30	15	0·5000		
3	30	23	0·7667		

Locate 1·567 on left-hand ordinate Σp scale, and 1·217 on the lower or upper abscissal Σpx scale. A straight edge, preferably thin and transparent, should be used as an aid in locating the position of the co-ordinate values. The estimates $\hat{\gamma}$, $\hat{\beta}$ can generally be read exactly to two significant figures with the third significant figure estimated by interpolation. In some sections of the nomograms an additional place of accuracy can be estimated. From the nomograms are found

$$\hat{\gamma} = 2{\cdot}008, \quad \hat{\beta} = 1{\cdot}30.$$

The values obtained by iterative computation are

$$\hat{\gamma} = 2{\cdot}0030, \quad \hat{\beta} = 1{\cdot}2992.$$

From the tables

$$s(\hat{\gamma}) = 0{\cdot}98/\sqrt{30} = 0{\cdot}18, \quad s(\hat{\beta}) = 1{\cdot}3/\sqrt{30} = 0{\cdot}24$$

so that

$$\hat{\gamma} = 2{\cdot}01 \pm 0{\cdot}18, \quad \hat{\beta} = 1{\cdot}30 \pm 0{\cdot}24.$$

Example 2 ($k = 4$ doses).

X	n	r	p		
0	30	10	0·3333	$\Sigma p = 2{\cdot}2334,$	$\Sigma px = \Sigma pX - 1{\cdot}5\Sigma p = 0{\cdot}7501.$
1	30	14	0·4667	$\Sigma pX = 4{\cdot}1002,$	
2	30	20	0·6667		
3	30	23	0·7667		

Locate 2·233 on the right-hand Σp scale and 0·750 on the abscissal Σpx scale. From the nomograms are found

$$\gamma = 1{\cdot}912, \quad \beta = 0{\cdot}652.$$

Since Σp was read on the right-hand scale,

$$\hat{\gamma} = 3 - \gamma = 1{\cdot}088.$$

The values obtained by iterative computation are

$$\hat{\gamma} = 1{\cdot}0935, \quad \hat{\beta} = 0{\cdot}6520.$$

From the tables, corresponding to $\gamma = 1{\cdot}91$ and $\beta = 0{\cdot}65$

$$s(\hat{\gamma}) = 1{\cdot}7/\sqrt{30} = 0{\cdot}31, \quad s(\hat{\beta}) = 1{\cdot}0/\sqrt{30} = 0{\cdot}18,$$

so that

$$\hat{\gamma} = 1{\cdot}09 \pm 0{\cdot}31, \quad \hat{\beta} = 0{\cdot}65 \pm 0{\cdot}18.$$

Three doses. Standard error of estimates, $s\sqrt{n}$

	γ = 1·0		1·1		1·2		1·3		1·4		1·45		1·5		1·55		1·6		1·65		1·7		1·75		1·8	
β	$s(\frac{\alpha}{\gamma})$	$s(\beta)$	$s(\frac{\alpha}{\gamma})$	$s(\beta)$	$s(\frac{\alpha}{\gamma})$	$s(\beta)$	$s(\frac{\alpha}{\gamma})$	$s(\beta)$	$s(\frac{\alpha}{\gamma})$	$s(\beta)$	$s(\frac{\alpha}{\gamma})$	$s(\beta)$	$s(\frac{\alpha}{\gamma})$	$s(\beta)$	$s(\frac{\alpha}{\gamma})$	$s(\beta)$	$s(\frac{\alpha}{\gamma})$	$s(\beta)$	$s(\frac{\alpha}{\gamma})$	$s(\beta)$	$s(\frac{\alpha}{\gamma})$	$s(\beta)$	$s(\frac{\alpha}{\gamma})$	$s(\beta)$	$s(\frac{\alpha}{\gamma})$	$s(\beta)$
0·50	2·4	1·5	2·4	1·5	2·4	1·5	2·5	1·5	2·6	1·5	2·7	1·5	2·7	1·5	2·8	1·5	2·9	1·5	3·0	1·5	3·1	1·5	3·2	1·5	3·2	1·5
·52	2·3	1·5	2·3	1·5	2·3	1·5	2·4	1·5	2·5	1·5	2·6	1·5	2·6	1·5	2·7	1·5	2·8	1·5	2·9	1·5	3·0	1·5	3·0	1·5	3·1	1·5
·54	2·2	1·5	2·2	1·5	2·2	1·5	2·3	1·5	2·4	1·5	2·5	1·5	2·5	1·5	2·6	1·5	2·7	1·5	2·8	1·5	2·9	1·5	2·9	1·5	3·0	1·5
·56	2·1	1·5	2·1	1·5	2·2	1·5	2·2	1·5	2·3	1·5	2·4	1·5	2·5	1·5	2·5	1·5	2·6	1·5	2·7	1·5	2·8	1·5	2·9	1·5	2·9	1·5
·58	2·0	1·5	2·0	1·5	2·1	1·5	2·2	1·5	2·3	1·5	2·3	1·5	2·4	1·5	2·4	1·5	2·5	1·5	2·6	1·5	2·7	1·5	2·7	1·5	2·8	1·5
0·60	2·0	1·5	2·0	1·5	2·0	1·5	2·1	1·5	2·2	1·5	2·2	1·5	2·3	1·5	2·3	1·5	2·4	1·5	2·5	1·5	2·6	1·5	2·6	1·5	2·7	1·5
·62	1·9	1·5	1·9	1·5	2·0	1·5	2·0	1·5	2·1	1·5	2·1	1·5	2·2	1·5	2·2	1·5	2·3	1·5	2·4	1·5	2·5	1·5	2·5	1·5	2·6	1·5
·64	1·9	1·5	1·9	1·5	1·9	1·5	2·0	1·5	2·0	1·5	2·1	1·5	2·1	1·5	2·2	1·5	2·3	1·5	2·3	1·5	2·4	1·5	2·5	1·5	2·5	1·5
·66	1·8	1·5	1·8	1·5	1·9	1·5	1·9	1·5	2·0	1·5	2·0	1·5	2·1	1·5	2·1	1·5	2·2	1·5	2·3	1·5	2·3	1·5	2·4	1·5	2·5	1·5
·68	1·8	1·5	1·8	1·5	1·8	1·5	1·9	1·5	1·9	1·5	2·0	1·5	2·0	1·5	2·1	1·5	2·1	1·5	2·2	1·5	2·3	1·5	2·3	1·5	2·4	1·5
0·70	1·7	1·5	1·7	1·5	1·8	1·5	1·8	1·5	1·9	1·5	1·9	1·5	2·0	1·5	2·0	1·5	2·1	1·5	2·1	1·5	2·2	1·5	2·3	1·6	2·3	1·6
·72	1·7	1·5	1·7	1·5	1·7	1·5	1·8	1·5	1·8	1·5	1·9	1·5	1·9	1·5	2·0	1·5	2·0	1·5	2·1	1·5	2·1	1·6	2·2	1·6	2·3	1·6
·74	1·6	1·5	1·6	1·5	1·7	1·5	1·7	1·5	1·8	1·5	1·8	1·5	1·9	1·5	1·9	1·5	2·0	1·5	2·0	1·6	2·1	1·6	2·1	1·6	2·2	1·6
·76	1·6	1·5	1·6	1·5	1·6	1·5	1·7	1·5	1·7	1·5	1·8	1·5	1·8	1·5	1·9	1·5	1·9	1·6	2·0	1·6	2·0	1·6	2·1	1·6	2·2	1·6
·78	1·6	1·5	1·6	1·5	1·6	1·5	1·6	1·5	1·7	1·5	1·7	1·5	1·8	1·6	1·8	1·6	1·9	1·6	1·9	1·6	2·0	1·6	2·0	1·6	2·1	1·6
0·80	1·5	1·5	1·5	1·5	1·6	1·5	1·6	1·5	1·7	1·5	1·7	1·6	1·7	1·6	1·8	1·6	1·8	1·6	1·9	1·6	1·9	1·6	2·0	1·6	2·0	1·6
·85	1·4	1·6	1·5	1·6	1·5	1·6	1·5	1·6	1·6	1·6	1·6	1·6	1·6	1·6	1·7	1·6	1·7	1·6	1·8	1·6	1·8	1·6	1·9	1·6	1·9	1·6
·90	1·4	1·6	1·4	1·6	1·4	1·6	1·4	1·6	1·5	1·6	1·5	1·6	1·5	1·6	1·6	1·6	1·6	1·6	1·7	1·6	1·7	1·6	1·8	1·6	1·8	1·7
·95	1·3	1·6	1·3	1·6	1·3	1·6	1·4	1·6	1·4	1·6	1·4	1·6	1·5	1·6	1·5	1·6	1·5	1·6	1·6	1·6	1·6	1·7	1·7	1·7	1·7	1·7
1·00	1·2	1·6	1·2	1·6	1·3	1·6	1·3	1·6	1·3	1·6	1·4	1·6	1·4	1·6	1·4	1·6	1·5	1·7	1·5	1·7	1·6	1·7	1·6	1·7	1·6	1·7
1·05	1·2	1·6	1·2	1·6	1·2	1·6	1·3	1·6	1·3	1·6	1·3	1·7	1·3	1·7	1·4	1·7	1·4	1·7	1·4	1·7	1·5	1·7	1·5	1·7	1·6	1·7
1·10	1·2	1·6	1·2	1·6	1·2	1·6	1·2	1·7	1·2	1·7	1·3	1·7	1·3	1·7	1·3	1·7	1·3	1·7	1·4	1·7	1·4	1·7	1·5	1·8	1·5	1·8
1·15	1·1	1·7	1·1	1·7	1·1	1·7	1·2	1·7	1·2	1·7	1·2	1·7	1·2	1·7	1·3	1·7	1·3	1·7	1·3	1·8	1·4	1·8	1·4	1·8	1·4	1·8
1·20	1·1	1·7	1·1	1·7	1·1	1·7	1·1	1·7	1·1	1·7	1·2	1·7	1·2	1·7	1·2	1·8	1·2	1·8	1·3	1·8	1·3	1·8	1·3	1·8	1·4	1·8
1·25	1·0	1·7	1·0	1·7	1·1	1·7	1·1	1·7	1·1	1·7	1·1	1·8	1·1	1·8	1·2	1·8	1·2	1·8	1·2	1·8	1·3	1·8	1·3	1·9	1·3	1·9
1·30	1·0	1·7	1·0	1·7	1·0	1·7	1·0	1·7	1·1	1·8	1·1	1·8	1·1	1·8	1·1	1·8	1·2	1·8	1·2	1·8	1·2	1·9	1·2	1·9	1·3	1·9
1·32	0·99	1·7	0·99	1·7	1·0	1·7	1·0	1·8	1·1	1·8	1·1	1·8	1·1	1·8	1·1	1·8	1·1	1·8	1·2	1·9	1·2	1·9	1·2	1·9	1·3	1·9
1·34	·98	1·7	·98	1·7	0·99	1·8	1·0	1·8	1·0	1·8	1·1	1·8	1·1	1·8	1·1	1·8	1·1	1·9	1·1	1·9	1·2	1·9	1·2	1·9	1·2	1·9
1·36	·97	1·8	·97	1·8	·98	1·8	0·99	1·8	1·0	1·8	1·0	1·8	1·1	1·8	1·1	1·8	1·1	1·9	1·1	1·9	1·2	1·9	1·2	1·9	1·2	2·0
1·38	·96	1·8	·96	1·8	·97	1·8	·99	1·8	1·0	1·8	1·0	1·8	1·0	1·8	1·1	1·9	1·1	1·9	1·1	1·9	1·1	1·9	1·2	1·9	1·2	2·0
1·40	0·95	1·8	0·95	1·8	0·96	1·8	0·98	1·8	1·0	1·8	1·0	1·8	1·0	1·9	1·1	1·9	1·1	1·9	1·1	1·9	1·1	1·9	1·2	2·0	1·2	2·0
1·42	·94	1·8	·94	1·8	·95	1·8	·97	1·8	0·99	1·8	1·0	1·9	1·0	1·9	1·0	1·9	1·1	1·9	1·1	1·9	1·1	2·0	1·1	2·0	1·2	2·0
1·44	·93	1·8	·93	1·8	·94	1·8	·96	1·8	·98	1·9	0·99	1·9	1·0	1·9	1·0	1·9	1·0	1·9	1·1	1·9	1·1	2·0	1·1	2·0	1·2	2·0
1·46	·92	1·8	·92	1·8	·93	1·8	·95	1·8	·97	1·9	·98	1·9	1·0	1·9	1·0	1·9	1·0	1·9	1·1	2·0	1·1	2·0	1·1	2·0	1·1	2·0
1·48	·91	1·8	·91	1·8	·92	1·8	·94	1·9	·96	1·9	·97	1·9	0·99	1·9	1·0	1·9	1·0	1·9	1·0	2·0	1·1	2·0	1·1	2·0	1·1	2·1
1·50	0·90	1·8	0·90	1·8	0·91	1·8	0·93	1·9	0·95	1·9	0·96	1·9	0·98	1·9	0·99	1·9	1·0	2·0	1·0	2·0	1·1	2·0	1·1	2·1	1·1	2·1
1·52	·89	1·8	·89	1·8	·90	1·9	·92	1·9	·94	1·9	·95	1·9	·97	1·9	·98	2·0	1·0	2·0	1·0	2·0	1·0	2·0	1·1	2·1	1·1	2·1
1·54	·88	1·9	·89	1·9	·89	1·9	·91	1·9	·93	1·9	·94	1·9	·96	1·9	·97	2·0	0·99	2·0	1·0	2·0	1·0	2·0	1·1	2·1	1·1	2·1
1·56	·87	1·9	·88	1·9	·89	1·9	·90	1·9	·92	1·9	·93	1·9	·95	2·0	·96	2·0	·98	2·0	1·0	2·0	1·0	2·1	1·0	2·1	1·1	2·1
1·58	·87	1·9	·87	1·9	·88	1·9	·89	1·9	·91	1·9	·92	2·0	·93	2·0	·95	2·0	·97	2·0	0·99	2·0	1·0	2·1	1·0	2·1	1·1	2·1
1·60	0·86	1·9	0·86	1·9	0·87	1·9	0·88	1·9	0·90	2·0	0·91	2·0	0·92	2·0	0·94	2·0	0·96	2·0	0·98	2·1	1·0	2·1	1·0	2·1	1·1	2·2
1·62	·85	1·9	·85	1·9	·86	1·9	·87	1·9	·89	2·0	·90	2·0	·92	2·0	·93	2·0	·95	2·1	·97	2·1	0·99	2·1	1·0	2·1	1·0	2·2
1·64	·84	1·9	·85	1·9	·85	1·9	·86	2·0	·88	2·0	·90	2·0	·91	2·0	·92	2·0	·94	2·1	·96	2·1	·98	2·1	0·99	2·2	1·0	2·2
1·66	·84	1·9	·84	1·9	·85	1·9	·86	2·0	·87	2·0	·89	2·0	·90	2·0	·91	2·1	·93	2·1	·95	2·1	·97	2·1	·99	2·2	1·0	2·2
1·68	·83	1·9	·83	1·9	·84	2·0	·85	2·0	·87	2·0	·88	2·0	·89	2·0	·90	2·1	·92	2·1	·94	2·1	·96	2·2	·98	2·2	1·0	2·2

Three doses (cont.). Standard error of estimates, $s\sqrt{n}$

Three doses (cont.). Standard error of estimates, $s\sqrt{n}$

	γ = 1·0		1·1		1·2		1·3		1·4		1·45		1·5		1·55		1·6		1·65		1·7		1·75		1·8	
β	$s(\frac{x}{\gamma})$	$s(\beta)$	$s(\frac{x}{\gamma})$	$s(\beta)$	$s(\frac{x}{\gamma})$	$s(\beta)$	$s(\frac{x}{\gamma})$	$s(\beta)$	$s(\frac{x}{\gamma})$	$s(\beta)$	$s(\frac{x}{\gamma})$	$s(\beta)$	$s(\frac{x}{\gamma})$	$s(\beta)$	$s(\frac{x}{\gamma})$	$s(\beta)$	$s(\frac{x}{\gamma})$	$s(\beta)$	$s(\frac{x}{\gamma})$	$s(\beta)$	$s(\frac{x}{\gamma})$	$s(\beta)$	$s(\frac{x}{\gamma})$	$s(\beta)$	$s(\frac{x}{\gamma})$	$s(\beta)$
1·70	0·82	2·0	0·82	2·0	0·83	2·0	0·84	2·0	0·86	2·0	0·87	2·0	0·88	2·1	0·89	2·1	0·91	2·1	0·93	2·1	0·95	2·2	0·97	2·2	0·99	2·3
1·72	·82	2·0	·82	2·0	·82	2·0	·83	2·0	·85	2·0	·86	2·1	·87	2·1	·88	2·1	·90	2·1	·92	2·2	·94	2·2	·96	2·2	·98	2·3
1·74	·81	2·0	·81	2·0	·82	2·0	·83	2·0	·84	2·1	·85	2·1	·86	2·1	·88	2·1	·89	2·1	·91	2·2	·93	2·2	·95	2·3	·97	2·3
1·76	·80	2·0	·81	2·0	·81	2·0	·82	2·0	·83	2·1	·84	2·1	·86	2·1	·87	2·1	·88	2·2	·90	2·2	·92	2·2	·94	2·3	·96	2·3
1·78	·80	2·0	·80	2·0	·80	2·0	·82	2·0	·83	2·1	·84	2·1	·85	2·1	·86	2·1	·87	2·2	·89	2·2	·91	2·2	·93	2·3	·95	2·3
1·80	0·79	2·0	0·79	2·0	0·80	2·0	0·81	2·1	0·82	2·1	0·83	2·1	0·84	2·1	0·85	2·2	0·87	2·2	0·88	2·2	0·90	2·3	0·92	2·3	0·94	2·4
1·82	·78	2·0	·79	2·0	·79	2·1	·80	2·1	·81	2·1	·82	2·1	·83	2·2	·85	2·2	·86	2·2	·87	2·2	·89	2·3	·91	2·4	·93	2·4
1·84	·78	2·1	·78	2·1	·79	2·1	·80	2·1	·81	2·1	·82	2·1	·83	2·2	·84	2·2	·85	2·2	·87	2·3	·88	2·3	·90	2·4	·93	2·4
1·86	·77	2·1	·78	2·1	·78	2·1	·79	2·1	·80	2·1	·81	2·2	·82	2·2	·83	2·2	·84	2·2	·86	2·3	·88	2·3	·90	2·4	·92	2·4
1·88	·77	2·1	·77	2·1	·77	2·1	·79	2·1	·80	2·2	·80	2·2	·81	2·2	·82	2·2	·84	2·3	·85	2·3	·87	2·3	·89	2·4	·91	2·4
1·90	0·76	2·1	0·76	2·1	0·77	2·1	0·78	2·1	0·79	2·2	0·80	2·2	0·81	2·2	0·82	2·2	0·83	2·3	0·84	2·3	0·86	2·4	0·88	2·4	0·90	2·5
1·92	·76	2·1	·76	2·1	·76	2·1	·77	2·2	·78	2·2	·79	2·2	·80	2·2	·81	2·3	·82	2·3	·84	2·3	·85	2·4	·87	2·4	·89	2·5
1·94	·75	2·1	·75	2·1	·76	2·1	·77	2·2	·78	2·2	·79	2·2	·79	2·2	·80	2·3	·81	2·3	·83	2·3	·84	2·4	·86	2·4	·88	2·5
1·96	·75	2·1	·75	2·2	·75	2·2	·76	2·2	·77	2·2	·78	2·2	·79	2·3	·80	2·3	·81	2·3	·82	2·4	·84	2·4	·86	2·5	·87	2·5
1·98	·74	2·2	·74	2·2	·75	2·2	·76	2·2	·77	2·2	·77	2·3	·78	2·3	·79	2·3	·80	2·3	·81	2·4	·83	2·4	·85	2·5	·87	2·5
2·00	0·74	2·2	0·74	2·2	0·74	2·2	0·75	2·2	0·76	2·3	0·77	2·3	0·78	2·3	0·79	2·3	0·80	2·4	0·81	2·4	0·82	2·4	0·84	2·5	0·86	2·6
2·02	·73	2·2	·73	2·2	·74	2·2	·75	2·2	·76	2·3	·76	2·3	·77	2·3	·78	2·3	·79	2·4	·80	2·4	·81	2·5	·83	2·5	·85	2·6
2·04	·73	2·2	·73	2·2	·73	2·2	·74	2·3	·75	2·3	·76	2·3	·76	2·3	·77	2·4	·78	2·4	·79	2·4	·81	2·5	·82	2·5	·84	2·6
2·06	·72	2·2	·72	2·2	·73	2·2	·74	2·3	·75	2·3	·75	2·3	·76	2·3	·77	2·4	·78	2·4	·79	2·5	·80	2·5	·82	2·6	·84	2·6
2·08	·72	2·3	·72	2·3	·72	2·3	·73	2·3	·74	2·3	·75	2·3	·75	2·4	·76	2·4	·77	2·4	·78	2·5	·79	2·5	·81	2·6	·83	2·7
2·10	0·71	2·3	0·72	2·3	0·72	2·3	0·73	2·3	0·74	2·3	0·74	2·4	0·75	2·4	0·76	2·4	0·77	2·5	0·78	2·5	0·79	2·5	0·80	2·6	0·82	2·7
2·12	·71	2·3	·71	2·3	·72	2·3	·72	2·3	·73	2·3	·74	2·4	·74	2·4	·75	2·4	·76	2·5	·77	2·5	·78	2·6	·80	2·6	·81	2·7
2·14	·71	2·3	·71	2·3	·71	2·3	·72	2·3	·73	2·4	·73	2·4	·74	2·4	·75	2·5	·75	2·5	·76	2·5	·78	2·6	·79	2·6	·81	2·7
2·16	·70	2·3	·70	2·3	·71	2·3	·71	2·4	·72	2·4	·73	2·4	·73	2·4	·74	2·5	·74	2·5	·75	2·6	·77	2·6	·78	2·7	·80	2·7
2·18	·70	2·3	·70	2·4	·70	2·4	·71	2·4	·72	2·4	·72	2·4	·73	2·4	·74	2·5	·74	2·5	·75	2·6	·76	2·6	·78	2·7	·79	2·8
2·20	0·69	2·4	0·69	2·4	0·70	2·4	0·71	2·4	0·71	2·4	0·72	2·4	0·72	2·5	0·73	2·5	0·74	2·5	0·75	2·6	0·76	2·6	0·77	2·7	0·79	2·8
2·22	·69	2·4	·69	2·4	·69	2·4	·70	2·4	·71	2·4	·71	2·5	·72	2·5	·72	2·5	·73	2·6	·74	2·6	·75	2·7	·76	2·7	·78	2·8
2·24	·69	2·4	·69	2·4	·69	2·4	·70	2·4	·71	2·5	·71	2·5	·72	2·5	·72	2·5	·73	2·6	·74	2·6	·75	2·7	·76	2·8	·78	2·8
2·26	·68	2·4	·68	2·4	·69	2·4	·69	2·4	·70	2·5	·71	2·5	·71	2·5	·72	2·6	·73	2·6	·74	2·6	·74	2·7	·76	2·8	·77	2·8
2·28	·68	2·4	·68	2·5	·68	2·5	·69	2·5	·70	2·5	·70	2·5	·71	2·5	·71	2·6	·72	2·6	·73	2·7	·74	2·7	·75	2·8	·76	2·9
2·30	0·67	2·5	0·68	2·5	0·68	2·5	0·69	2·5	0·69	2·5	0·70	2·5	0·70	2·6	0·71	2·6	0·72	2·6	0·72	2·7	0·73	2·7	0·74	2·8	0·76	2·9
2·35	·67	2·5	·67	2·5	·67	2·5	·68	2·5	·69	2·5	·69	2·6	·69	2·6	·70	2·6	·71	2·7	·71	2·7	·72	2·8	·73	2·9	·75	2·9
2·40	·66	2·6	·66	2·6	·66	2·6	·67	2·6	·68	2·6	·68	2·6	·68	2·7	·69	2·7	·69	2·7	·70	2·8	·71	2·8	·72	2·9	·73	3·0
2·45	·65	2·6	·65	2·6	·65	2·6	·66	2·6	·67	2·6	·67	2·7	·67	2·7	·68	2·7	·68	2·8	·69	2·8	·70	2·9	·71	3·0	·72	3·1
2·50	·64	2·7	·64	2·7	·65	2·7	·65	2·7	·66	2·7	·66	2·7	·66	2·7	·67	2·8	·67	2·8	·68	2·9	·69	2·9	·69	3·0	·70	3·1
2·55	·63	2·7	·63	2·7	·64	2·7	·64	2·7	·65	2·7	·65	2·8	·66	2·8	·66	2·8	·67	2·9	·67	2·9	·68	3·0	·68	3·1	·69	3·2
2·60	·62	2·8	·63	2·8	·63	2·8	·64	2·8	·64	2·8	·65	2·8	·65	2·8	·65	2·9	·66	2·9	·66	3·0	·67	3·0	·67	3·1	·68	3·2
2·65	·62	2·8	·62	2·8	·62	2·8	·63	2·8	·64	2·8	·64	2·8	·64	2·9	·65	2·9	·65	3·0	·65	3·0	·66	3·1	·66	3·2	·67	3·3
2·70	·61	2·9	·61	2·9	·62	2·9	·62	2·9	·63	2·9	·63	2·9	·64	2·9	·64	3·0	·64	3·0	·65	3·1	·65	3·2	·65	3·3	·66	3·4
2·75	·60	3·0	·60	3·0	·61	2·9	·62	2·9	·62	2·9	·62	2·9	·63	3·0	·63	3·0	·63	3·0	·64	3·1	·64	3·2	·65	3·3	·65	3·4
2·80	·60	3·0	·60	3·0	·60	3·0	·61	3·0	·62	3·0	·62	3·0	·62	3·0	·62	3·0	·63	3·1	·63	3·2	·63	3·3	·64	3·4	·64	3·5
2·85	·59	3·1	·60	3·1	·60	3·1	·60	3·0	·61	3·0	·61	3·0	·61	3·1	·61	3·1	·62	3·1	·62	3·2	·62	3·3	·63	3·4	·63	3·6
2·90	·58	3·2	·59	3·2	·59	3·1	·60	3·1	·61	3·1	·61	3·1	·61	3·1	·61	3·1	·61	3·2	·61	3·2	·61	3·3	·61	3·4	·62	3·5	·62	3·6
2·95	·58	3·3	·58	3·2	·59	3·2	·60	3·1	·60	3·1	·60	3·1	·61	3·1	·61	3·2	·61	3·2	·61	3·3	·61	3·4	·61	3·5	·61	3·7
3·0	0·57	3·3	0·57	3·3	0·58	3·2	0·59	3·2	0·60	3·2	0·60	3·2	0·60	3·2	0·60	3·2	0·60	3·3	0·60	3·4	0·60	3·5	0·60	3·6	0·61	3·8
3·1	·56	3·5	·56	3·5	·57	3·4	·58	3·3	·59	3·3	·59	3·3	·59	3·3	·59	3·3	·59	3·4	·59	3·5	·59	3·6	·59	3·7	·59	3·9
3·2	·55	3·6	·55	3·6	·56	3·5	·57	3·4	·58	3·4	·58	3·4	·58	3·4	·58	3·4	·58	3·5	·58	3·6	·58	3·7	·58	3·9	·58	4·0
3·3	·54	3·8	·54	3·8	·55	3·7	·56	3·6	·57	3·5	·57	3·5	·58	3·5	·58	3·5	·57	3·6	·57	3·7	·57	3·8	·57	4·0	·56	4·2
3·4	·53	4·0	·53	3·9	·54	3·8	·55	3·7	·56	3·6	·57	3·6	·57	3·6	·57	3·6	·57	3·7	·56	3·8	·56	3·9	·56	4·1	·55	4·3
3·5	·52	4·2	·52	4·1	·53	4·0	·54	3·8	·56	3·7	·56	3·7	·56	3·7	·56	3·7	·56	3·8	·56	3·9	·55	4·1	·55	4·3	·54	4·5

Four doses. Standard error of estimates, $s\sqrt{n}$

	γ = 1·5		1·6		1·7		1·8		1·9		2·0		2·1		2·2		2·25		2·3		2·35		2·4		2·45		2·5	
β	$s(\hat{\gamma})$	$s(\hat{\beta})$	$s(\hat{\gamma})$	$s(\hat{\beta})$	$s(\hat{\gamma})$	$s(\hat{\beta})$	$s(\hat{\gamma})$	$s(\hat{\beta})$	$s(\hat{\gamma})$	$s(\hat{\beta})$	$s(\hat{\gamma})$	$s(\hat{\beta})$	$s(\hat{\gamma})$	$s(\hat{\beta})$	$s(\hat{\gamma})$	$s(\hat{\beta})$	$s(\hat{\gamma})$	$s(\hat{\beta})$	$s(\hat{\gamma})$	$s(\hat{\beta})$	$s(\hat{\gamma})$	$s(\hat{\beta})$	$s(\hat{\gamma})$	$s(\hat{\beta})$	$s(\hat{\gamma})$	$s(\hat{\beta})$	$s(\hat{\gamma})$	$s(\hat{\beta})$
0·50	2·1	0·95	2·1	0·95	2·1	0·95	2·1	0·95	2·2	0·96	2·2	0·96	2·3	0·96	2·4	0·97	2·4	0·97	2·5	0·97	2·5	0·97	2·6	0·97	2·7	0·98	2·7	0·98
·52	2·0	·96	2·0	·96	2·0	·96	2·1	·96	2·1	·96	2·2	·96	2·2	·97	2·3	·97	2·4	·97	2·4	·98	2·5	·98	2·5	·98	2·6	·98	2·6	·99
·54	1·9	·96	1·9	·96	2·0	·96	2·0	·96	2·0	·97	2·1	·97	2·2	·97	2·2	·98	2·3	·98	2·3	·98	2·4	·98	2·4	·99	2·5	·99	2·5	·99
·56	1·9	·97	1·9	·97	1·9	·97	1·9	·97	2·0	·97	2·0	·98	2·1	·98	2·2	·98	2·2	·99	2·2	·99	2·3	·99	2·3	·99	2·4	1·0	2·4	1·0
·58	1·8	·97	1·8	·97	1·8	·97	1·9	·98	1·9	·98	2·0	·98	2·0	·99	2·1	·99	2·1	·99	2·2	1·0	2·2	1·0	2·3	1·0	2·3	1·0	2·3	1·0
0·60	1·8	0·98	1·8	0·98	1·8	0·98	1·8	0·98	1·8	0·98	1·9	0·99	1·9	0·99	2·0	0·99	2·1	1·0	2·1	1·0	2·1	1·0	2·2	1·0	2·2	1·0	2·3	1·0
·62	1·7	·98	1·7	·98	1·7	·98	1·8	·99	1·8	·99	1·8	·99	1·9	1·0	2·0	1·0	2·0	1·0	2·0	1·0	2·1	1·0	2·1	1·0	2·2	1·0	2·2	1·0
·64	1·7	·99	1·7	·99	1·7	·99	1·7	·99	1·7	1·0	1·8	1·0	1·8	1·0	1·9	1·0	1·9	1·0	2·0	1·0	2·0	1·0	2·0	1·0	2·1	1·0	2·1	1·0
·66	1·6	·99	1·6	1·0	1·6	1·0	1·7	1·0	1·7	1·0	1·7	1·0	1·8	1·0	1·8	1·0	1·9	1·0	1·9	1·0	1·9	1·0	2·0	1·0	2·0	1·0	2·1	1·0
·68	1·6	1·0	1·6	1·0	1·6	1·0	1·6	1·0	1·6	1·0	1·7	1·0	1·7	1·0	1·8	1·0	1·8	1·0	1·9	1·0	1·9	1·0	1·9	1·0	2·0	1·0	2·0	1·1
0·70	1·5	1·0	1·5	1·0	1·6	1·0	1·6	1·0	1·6	1·0	1·6	1·0	1·7	1·0	1·8	1·0	1·8	1·0	1·8	1·0	1·8	1·0	1·9	1·1	1·9	1·1	2·0	1·1
·72	1·5	1·0	1·5	1·0	1·5	1·0	1·5	1·0	1·6	1·0	1·6	1·0	1·6	1·0	1·7	1·0	1·7	1·0	1·8	1·0	1·8	1·1	1·8	1·1	1·9	1·1	1·9	1·1
·74	1·5	1·0	1·5	1·0	1·5	1·0	1·5	1·0	1·5	1·0	1·6	1·0	1·6	1·0	1·7	1·0	1·7	1·1	1·7	1·1	1·8	1·1	1·8	1·1	1·8	1·1	1·9	1·1
·76	1·4	1·0	1·4	1·0	1·4	1·0	1·5	1·0	1·5	1·0	1·5	1·0	1·6	1·0	1·6	1·1	1·6	1·1	1·7	1·1	1·7	1·1	1·7	1·1	1·8	1·1	1·8	1·1
·78	1·4	1·0	1·4	1·0	1·4	1·0	1·5	1·0	1·5	1·0	1·5	1·0	1·5	1·1	1·6	1·1	1·6	1·1	1·6	1·1	1·7	1·1	1·7	1·1	1·7	1·1	1·8	1·1
0·80	1·4	1·0	1·4	1·0	1·4	1·0	1·4	1·0	1·4	1·1	1·5	1·1	1·5	1·1	1·5	1·1	1·6	1·1	1·6	1·1	1·6	1·1	1·7	1·1	1·7	1·1	1·7	1·1
·82	1·3	1·0	1·3	1·1	1·4	1·1	1·4	1·1	1·4	1·1	1·4	1·1	1·5	1·1	1·5	1·1	1·5	1·1	1·6	1·1	1·6	1·1	1·6	1·1	1·7	1·1	1·7	1·1
·84	1·3	1·1	1·3	1·1	1·3	1·1	1·3	1·1	1·4	1·1	1·4	1·1	1·4	1·1	1·5	1·1	1·5	1·1	1·5	1·1	1·5	1·1	1·5	1·1	1·6	1·1	1·6	1·1
·86	1·3	1·1	1·3	1·1	1·3	1·1	1·3	1·1	1·3	1·1	1·4	1·1	1·4	1·1	1·4	1·1	1·5	1·1	1·5	1·1	1·5	1·1	1·5	1·1	1·6	1·1	1·6	1·1
·88	1·3	1·1	1·3	1·1	1·3	1·1	1·3	1·1	1·3	1·1	1·3	1·1	1·4	1·1	1·4	1·1	1·4	1·1	1·5	1·1	1·5	1·1	1·5	1·1	1·5	1·1	1·6	1·2
0·90	1·2	1·1	1·2	1·1	1·3	1·1	1·3	1·1	1·3	1·1	1·3	1·1	1·3	1·1	1·4	1·1	1·4	1·1	1·4	1·1	1·4	1·1	1·5	1·2	1·5	1·2	1·5	1·2
1·00	1·1	1·1	1·1	1·1	1·2	1·1	1·2	1·1	1·2	1·1	1·2	1·2	1·2	1·2	1·3	1·2	1·3	1·2	1·3	1·2	1·3	1·2	1·4	1·2	1·4	1·2	1·4	1·2
1·10	1·1	1·2	1·1	1·2	1·1	1·2	1·1	1·2	1·1	1·2	1·1	1·2	1·1	1·2	1·2	1·2	1·2	1·2	1·2	1·3	1·2	1·3	1·2	1·3	1·3	1·3	1·3	1·3
1·20	1·0	1·2	1·0	1·2	1·0	1·2	1·0	1·2	1·0	1·3	1·0	1·3	1·1	1·3	1·1	1·3	1·1	1·3	1·1	1·3	1·1	1·3	1·1	1·3	1·2	1·4	1·2	1·4
1·22	0·98	1·2	0·99	1·2	0·99	1·2	1·0	1·3	1·0	1·3	1·0	1·3	1·0	1·3	1·1	1·3	1·1	1·3	1·1	1·3	1·1	1·3	1·1	1·4	1·1	1·4	1·2	1·4
1·24	·97	1·3	·97	1·3	·98	1·3	0·99	1·3	1·0	1·3	1·0	1·3	1·0	1·3	1·0	1·3	1·1	1·3	1·1	1·3	1·1	1·4	1·1	1·4	1·1	1·4	1·2	1·4
1·26	·96	1·3	·96	1·3	·97	1·3	·98	1·3	0·99	1·3	1·0	1·3	1·0	1·3	1·0	1·3	1·0	1·3	1·1	1·4	1·1	1·4	1·1	1·4	1·1	1·4	1·1	1·4
1·28	·95	1·3	·95	1·3	·96	1·3	·97	1·3	·98	1·3	0·99	1·3	1·0	1·3	1·0	1·4	1·0	1·4	1·0	1·4	1·1	1·4	1·1	1·4	1·1	1·4	1·1	1·4
1·30	0·94	1·3	0·94	1·3	0·95	1·3	0·95	1·3	0·96	1·3	0·98	1·3	0·99	1·3	1·0	1·4	1·0	1·4	1·0	1·4	1·0	1·4	1·1	1·4	1·1	1·4	1·1	1·5
1·32	·93	1·3	·93	1·3	·94	1·3	·94	1·3	·95	1·3	·96	1·3	·98	1·4	0·99	1·4	1·0	1·4	1·0	1·4	1·0	1·4	1·1	1·4	1·1	1·5	1·1	1·5
1·34	·92	1·3	·92	1·3	·93	1·3	·93	1·3	·94	1·3	·95	1·4	·97	1·4	·98	1·4	0·99	1·4	1·0	1·4	1·0	1·4	1·0	1·4	1·1	1·5	1·1	1·5
1·36	·91	1·3	·91	1·3	·92	1·3	·92	1·3	·93	1·4	·94	1·4	·96	1·4	·97	1·4	·99	1·4	1·0	1·4	1·0	1·4	1·0	1·5	1·0	1·5	1·1	1·5
1·38	·90	1·3	·90	1·3	·91	1·3	·91	1·3	·92	1·4	·93	1·4	·95	1·4	·96	1·4	·98	1·4	0·99	1·4	1·0	1·5	1·0	1·5	1·0	1·5	1·0	1·5
1·40	0·89	1·4	0·90	1·4	0·90	1·4	0·90	1·4	0·91	1·4	0·92	1·4	0·93	1·4	0·95	1·4	0·96	1·5	0·98	1·5	0·99	1·5	1·0	1·5	1·0	1·5	1·0	1·5
1·42	·89	1·4	·89	1·4	·89	1·4	·90	1·4	·90	1·4	·91	1·4	·93	1·4	·94	1·5	·95	1·5	·96	1·5	·98	1·5	0·99	1·5	1·0	1·5	1·0	1·6
1·44	·88	1·4	·88	1·4	·88	1·4	·89	1·4	·89	1·4	·90	1·4	·92	1·4	·93	1·5	·94	1·5	·95	1·5	·97	1·5	·98	1·5	0·99	1·5	1·0	1·6
1·46	·87	1·4	·87	1·4	·87	1·4	·88	1·4	·89	1·4	·90	1·4	·91	1·5	·92	1·5	·93	1·5	·94	1·5	·96	1·5	·97	1·5	·98	1·6	1·0	1·6
1·48	·86	1·4	·86	1·4	·87	1·4	·87	1·4	·88	1·4	·89	1·5	·90	1·5	·91	1·5	·92	1·5	·93	1·5	·94	1·5	·96	1·6	·98	1·6	0·99	1·6
1·50	0·86	1·4	0·86	1·4	0·86	1·4	0·86	1·4	0·87	1·4	0·88	1·5	0·89	1·5	0·90	1·5	0·91	1·5	0·92	1·5	0·93	1·6	0·95	1·6	0·96	1·6	0·98	1·6
1·52	·85	1·4	·85	1·4	·85	1·4	·85	1·4	·86	1·5	·87	1·5	·88	1·5	·89	1·5	·90	1·5	·91	1·6	·92	1·6	·94	1·6	·95	1·6	·97	1·6
1·54	·84	1·4	·84	1·4	·84	1·5	·85	1·5	·85	1·5	·86	1·5	·87	1·5	·88	1·5	·89	1·6	·90	1·6	·92	1·6	·93	1·6	·94	1·6	·95	1·7
1·56	·83	1·5	·84	1·5	·84	1·5	·84	1·5	·85	1·5	·86	1·5	·86	1·5	·87	1·6	·88	1·6	·90	1·6	·91	1·6	·92	1·6	·93	1·7	·94	1·7
1·58	·83	1·5	·83	1·5	·83	1·5	·83	1·5	·84	1·5	·85	1·5	·86	1·6	·87	1·6	·88	1·6	·89	1·6	·90	1·6	·91	1·6	·92	1·7	·93	1·7

1·54	·84	1·4	·84	1·4	·84	1·5	·85	1·5	·85	1·5	·86	1·5	·87	1·5	·88	1·5	·89	1·6	·90	1·6	·92	1·6	·93	1·6	·94	1·6	·95	1·7
1·56	·83	1·5	·84	1·5	·84	1·5	·84	1·5	·85	1·5	·86	1·5	·86	1·5	·87	1·6	·88	1·6	·90	1·6	·91	1·6	·92	1·6	·93	1·7	·94	1·7
1·58	·83	1·5	·83	1·5	·83	1·5	·83	1·5	·84	1·5	·85	1·5	·86	1·6	·87	1·6	·88	1·6	·89	1·6	·90	1·6	·91	1·6	·92	1·7	·93	1·7

Four doses (cont.). Standard error of estimates, $s\sqrt{n}$

	γ																											
	1·5		1·6		1·7		1·8		1·9		2·0		2·1		2·2		2·25		2·3		2·35		2·4		2·45		2·5	
β	$s(\frac{\alpha}{\gamma})$	$s(\beta)$	$s(\frac{\alpha}{\gamma})$	$s(\beta)$	$s(\frac{\alpha}{\gamma})$	$s(\beta)$	$s(\frac{\alpha}{\gamma})$	$s(\beta)$	$s(\frac{\alpha}{\gamma})$	$s(\beta)$	$s(\frac{\alpha}{\gamma})$	$s(\beta)$	$s(\frac{\alpha}{\gamma})$	$s(\beta)$	$s(\frac{\alpha}{\gamma})$	$s(\beta)$	$s(\frac{\alpha}{\gamma})$	$s(\beta)$	$s(\frac{\alpha}{\gamma})$	$s(\beta)$	$s(\frac{\alpha}{\gamma})$	$s(\beta)$	$s(\frac{\alpha}{\gamma})$	$s(\beta)$	$s(\frac{\alpha}{\gamma})$	$s(\beta)$	$s(\frac{\alpha}{\gamma})$	$s(\beta)$
1·60	0·82	1·5	0·82	1·5	0·82	1·5	0·83	1·5	0·83	1·5	0·84	1·5	0·85	1·6	0·86	1·6	0·87	1·6	0·88	1·6	0·89	1·6	0·90	1·7	0·91	1·7	0·92	1·7
1·62	·81	1·5	·81	1·5	·82	1·5	·82	1·5	·83	1·5	·83	1·6	·84	1·6	·85	1·6	·86	1·6	·87	1·6	·88	1·7	·89	1·7	·90	1·7	·91	1·7
1·64	·81	1·5	·81	1·5	·81	1·5	·81	1·5	·82	1·6	·83	1·6	·83	1·6	·84	1·6	·85	1·6	·86	1·7	·87	1·7	·88	1·7	·89	1·7	·91	1·8
1·66	·80	1·5	·80	1·5	·80	1·5	·81	1·6	·81	1·6	·82	1·6	·83	1·6	·84	1·6	·85	1·7	·85	1·7	·86	1·7	·87	1·7	·88	1·7	·90	1·8
1·68	·80	1·6	·80	1·6	·80	1·6	·80	1·6	·81	1·6	·81	1·6	·82	1·6	·83	1·7	·84	1·7	·85	1·7	·86	1·7	·87	1·7	·88	1·8	·89	1·8
1·70	0·79	1·6	0·79	1·6	0·79	1·6	0·80	1·6	0·80	1·6	0·81	1·6	0·82	1·6	0·83	1·7	0·83	1·7	0·84	1·7	0·85	1·7	0·86	1·8	0·87	1·8	0·88	1·8
1·72	·78	1·6	·79	1·6	·79	1·6	·79	1·6	·79	1·6	·80	1·6	·81	1·7	·82	1·7	·83	1·7	·83	1·7	·84	1·8	·85	1·8	·86	1·8	·87	1·8
1·74	·78	1·6	·78	1·6	·78	1·6	·78	1·6	·79	1·6	·80	1·7	·80	1·7	·81	1·7	·82	1·7	·83	1·8	·83	1·8	·84	1·8	·85	1·8	·86	1·8
1·76	·77	1·6	·78	1·6	·78	1·6	·78	1·6	·78	1·7	·79	1·7	·80	1·7	·81	1·7	·81	1·8	·82	1·8	·83	1·8	·84	1·8	·85	1·8	·86	1·9
1·78	·77	1·6	·77	1·6	·77	1·6	·77	1·7	·78	1·7	·78	1·7	·79	1·7	·80	1·8	·81	1·8	·81	1·8	·82	1·8	·83	1·8	·84	1·9	·85	1·9
1·80	0·76	1·6	0·76	1·7	0·77	1·7	0·77	1·7	0·77	1·7	0·78	1·7	0·78	1·7	0·79	1·8	0·80	1·8	0·81	1·8	0·81	1·8	0·82	1·9	0·83	1·9	0·84	1·9
1·82	·76	1·7	·76	1·7	·76	1·7	·76	1·7	·77	1·7	·77	1·7	·78	1·8	·79	1·8	·79	1·8	·80	1·8	·81	1·8	·81	1·9	·82	1·9	·83	1·9
1·84	·75	1·7	·75	1·7	·76	1·7	·76	1·7	·76	1·7	·77	1·7	·77	1·8	·78	1·8	·79	1·8	·79	1·8	·80	1·9	·81	1·9	·82	1·9	·83	1·9
1·86	·75	1·7	·75	1·7	·76	1·7	·76	1·7	·76	1·7	·76	1·8	·77	1·8	·78	1·8	·78	1·8	·79	1·9	·79	1·9	·80	1·9	·81	1·9	·82	2·0
1·88	·74	1·7	·75	1·7	·75	1·7	·75	1·7	·75	1·8	·76	1·8	·76	1·8	·77	1·8	·78	1·9	·78	1·9	·79	1·9	·79	1·9	·80	2·0	·81	2·0
1·90	0·74	1·7	0·74	1·7	0·74	1·7	0·74	1·8	0·75	1·8	0·75	1·8	0·76	1·8	0·77	1·9	0·77	1·9	0·78	1·9	0·78	1·9	0·79	1·9	0·80	2·0	0·80	2·0
1·92	·74	1·7	·74	1·8	·74	1·8	·74	1·8	·74	1·8	·75	1·8	·75	1·8	·76	1·9	·76	1·9	·77	1·9	·78	2·0	·78	2·0	·79	2·0	·80	2·0
1·94	·73	1·8	·73	1·8	·73	1·8	·73	1·8	·74	1·8	·74	1·8	·75	1·9	·76	1·9	·76	1·9	·76	1·9	·77	2·0	·78	2·0	·78	2·0	·79	2·0
1·96	·73	1·8	·73	1·8	·73	1·8	·73	1·8	·73	1·8	·74	1·9	·74	1·9	·75	1·9	·75	1·9	·76	2·0	·77	2·0	·77	2·0	·78	2·0	·79	2·1
1·98	·72	1·8	·72	1·8	·72	1·8	·73	1·8	·73	1·9	·73	1·9	·74	1·9	·75	1·9	·75	2·0	·75	2·0	·76	2·0	·77	2·0	·77	2·1	·78	2·1
2·00	0·72	1·8	0·72	1·8	0·72	1·8	0·72	1·9	0·72	1·9	0·73	1·9	0·73	1·9	0·74	2·0	0·74	2·0	0·75	2·0	0·75	2·0	0·76	2·0	0·77	2·1	0·78	2·1
2·02	·72	1·8	·72	1·8	·72	1·9	·72	1·9	·72	1·9	·72	1·9	·73	1·9	·73	2·0	·74	2·0	·74	2·0	·75	2·0	·76	2·1	·76	2·1	·77	2·1
2·04	·71	1·9	·71	1·9	·71	1·9	·71	1·9	·72	1·9	·72	1·9	·72	2·0	·73	2·0	·73	2·0	·74	2·0	·74	2·1	·75	2·1	·76	2·1	·76	2·1
2·06	·71	1·9	·71	1·9	·71	1·9	·71	1·9	·71	1·9	·72	2·0	·72	2·0	·73	2·0	·73	2·0	·73	2·1	·74	2·1	·75	2·1	·75	2·1	·76	2·2
2·08	·70	1·9	·70	1·9	·70	1·9	·70	1·9	·71	2·0	·71	2·0	·71	2·0	·72	2·0	·73	2·1	·73	2·1	·73	2·1	·74	2·1	·75	2·2	·75	2·2
2·10	0·70	1·9	0·70	1·9	0·70	1·9	0·70	1·9	0·70	2·0	0·71	2·0	0·71	2·0	0·72	2·1	0·72	2·1	0·73	2·1	0·73	2·1	0·74	2·1	0·74	2·2	0·75	2·2
2·12	·70	1·9	·70	1·9	·70	1·9	·70	2·0	·70	2·0	·70	2·0	·71	2·0	·71	2·1	·72	2·1	·72	2·1	·73	2·1	·73	2·2	·74	2·2	·74	2·2
2·14	·69	2·0	·69	2·0	·69	2·0	·69	2·0	·70	2·0	·70	2·0	·70	2·1	·71	2·1	·71	2·1	·72	2·1	·72	2·2	·73	2·2	·73	2·2	·74	2·3
2·16	·69	2·0	·69	2·0	·69	2·0	·69	2·0	·69	2·0	·70	2·1	·70	2·1	·70	2·1	·71	2·1	·71	2·2	·72	2·2	·72	2·2	·73	2·2	·73	2·3
2·18	·69	2·0	·69	2·0	·69	2·0	·69	2·0	·69	2·1	·69	2·1	·69	2·1	·70	2·1	·70	2·1	·71	2·2	·71	2·2	·72	2·2	·72	2·3	·73	2·3
2·20	0·68	2·0	0·68	2·0	0·68	2·0	0·68	2·0	0·68	2·1	0·68	2·1	0·69	2·1	0·70	2·2	0·70	2·2	0·70	2·2	0·71	2·2	0·71	2·3	0·72	2·3	0·72	2·3
2·25	·67	2·1	·67	2·1	·67	2·1	·67	2·1	·68	2·1	·68	2·2	·68	2·2	·69	2·2	·69	2·2	·69	2·3	·70	2·3	·70	2·3	·71	2·3	·71	2·4
2·30	·67	2·1	·67	2·1	·67	2·1	·67	2·2	·67	2·2	·67	2·2	·67	2·2	·68	2·3	·68	2·3	·69	2·3	·69	2·3	·69	2·3	·70	2·4	·70	2·4
2·35	·66	2·2	·66	2·2	·66	2·2	·66	2·2	·66	2·2	·66	2·3	·66	2·3	·67	2·3	·67	2·3	·68	2·4	·68	2·4	·69	2·4	·69	2·4	·69	2·5
2·40	·65	2·2	·65	2·2	·65	2·2	·65	2·3	·65	2·3	·65	2·3	·65	2·4	·66	2·4	·67	2·4	·67	2·4	·67	2·4	·68	2·5	·68	2·5	·68	2·5
2·45	·64	2·3	·64	2·3	·64	2·3	·64	2·3	·64	2·4	·64	2·4	·65	2·4	·65	2·4	·66	2·5	·66	2·5	·66	2·5	·67	2·5	·67	2·5	·67	2·6
2·50	·64	2·3	·64	2·3	·64	2·3	·64	2·4	·64	2·4	·64	2·5	·64	2·5	·65	2·5	·65	2·5	·65	2·5	·65	2·5	·66	2·6	·66	2·6	·67	2·6
2·55	·63	2·4	·63	2·4	·63	2·4	·63	2·5	·63	2·5	·63	2·5	·64	2·6	·64	2·6	·64	2·6	·64	2·6	·65	2·6	·65	2·6	·65	2·6	·66	2·7
2·60	·63	2·4	·63	2·4	·63	2·5	·62	2·5	·62	2·5	·62	2·6	·62	2·6	·63	2·6	·64	2·6	·64	2·6	·64	2·6	·64	2·7	·64	2·7	·65	2·7
2·65	·62	2·5	·62	2·5	·62	2·5	·61	2·6	·61	2·6	·61	2·6	·62	2·7	·62	2·7	·63	2·7	·63	2·7	·63	2·7	·64	2·7	·64	2·7	·64	2·8
2·70	·62	2·5	·62	2·5	·61	2·6	·61	2·6	·61	2·7	·61	2·7	·61	2·7	·62	2·7	·62	2·8	·62	2·8	·63	2·8	·63	2·8	·63	2·8	·64	2·8
2·75	·61	2·6	·61	2·6	·61	2·6	·61	2·7	·60	2·7	·60	2·8	·60	2·8	·61	2·8	·61	2·8	·62	2·8	·62	2·8	·62	2·8	·63	2·9	·63	2·9
2·8	0·61	2·6	0·60	2·6	0·60	2·7	0·60	2·7	0·59	2·8	0·60	2·9	0·60	2·9	0·60	2·9	0·61	2·9	0·61	2·9	0·61	2·9	0·62	2·9	0·62	2·9	0·62	2·9
2·9	·60	2·7	·59	2·8	·59	2·8	·59	2·9	·58	3·0	·58	3·0	·59	3·0	·59	3·0	·60	3·0	·60	3·0	·60	3·0	·61	3·0	·61	3·0	·61	3·0
3·0	·59	2·8	·59	2·9	·58	2·9	·58	3·0	·57	3·1	·57	3·2	·57	3·2	·58	3·2	·59	3·1	·59	3·1	·59	3·1	·60	3·1	·60	3·1	·60	3·1

Five doses. Standard error of estimates, $s\sqrt{n}$

	γ																	
	2·0		2·1		2·2		2·3		2·4		2·5		2·6		2·7		2·8		2·9	
β	$s(\frac{\alpha}{\gamma})$	$s(\beta)$	$s(\frac{\alpha}{\gamma})$	$s(\beta)$	$s(\frac{\alpha}{\gamma})$	$s(\beta)$	$s(\frac{\alpha}{\gamma})$	$s(\beta)$	$s(\frac{\alpha}{\gamma})$	$s(\beta)$	$s(\frac{\alpha}{\gamma})$	$s(\beta)$	$s(\frac{\alpha}{\gamma})$	$s(\beta)$	$s(\frac{\alpha}{\gamma})$	$s(\beta)$	$s(\frac{\alpha}{\gamma})$	$s(\beta)$	$s(\frac{\alpha}{\gamma})$	$s(\beta)$
0·40	2·3	0·68	2·3	0·68	2·3	0·68	2·4	0·68	2·4	0·68	2·4	0·68	2·5	0·68	2·6	0·68	2·6	0·68	2·7	0·69
·42	2·2	·68	2·2	·68	2·2	·68	2·3	·68	2·3	·68	2·3	·68	2·4	·68	2·4	·69	2·5	·69	2·6	·69
·44	2·1	·68	2·1	·68	2·1	·68	2·2	·69	2·2	·69	2·2	·69	2·3	·69	2·3	·69	2·4	·69	2·5	·70
·46	2·0	·69	2·0	·69	2·1	·69	2·1	·69	2·1	·69	2·1	·69	2·2	·70	2·2	·70	2·3	·70	2·4	·70
·48	2·0	·69	2·0	·69	2·0	·70	2·0	·70	2·0	·70	2·1	·70	2·1	·70	2·2	·70	2·2	·71	2·3	·71
0·50	1·9	0·70	1·9	0·70	1·9	0·70	1·9	0·70	2·0	0·70	2·0	0·70	2·0	0·71	2·1	0·71	2·1	0·71	2·2	0·72
·52	1·8	·71	1·8	·71	1·8	·71	1·9	·71	1·9	·71	1·9	·71	2·0	·71	2·0	·72	2·0	·72	2·1	·72
·54	1·8	·71	1·8	·71	1·8	·71	1·8	·71	1·8	·71	1·8	·72	1·9	·72	1·9	·72	2·0	·72	2·0	·73
·56	1·7	·72	1·7	·72	1·7	·72	1·7	·72	1·8	·72	1·8	·72	1·8	·73	1·9	·73	1·9	·73	2·0	·74
·58	1·7	·72	1·7	·72	1·7	·72	1·7	·73	1·7	·73	1·7	·73	1·8	·73	1·8	·74	1·8	·74	1·9	·74
0·60	1·6	0·73	1·6	0·73	1·6	0·73	1·6	0·73	1·7	0·73	1·7	0·74	1·7	0·74	1·8	0·74	1·8	0·75	1·8	0·75
·62	1·6	·74	1·6	·74	1·6	·74	1·6	·74	1·6	·74	1·6	·74	1·7	·75	1·7	·75	1·7	·75	1·8	·76
·64	1·5	·74	1·5	·74	1·5	·74	1·6	·75	1·6	·75	1·6	·75	1·6	·75	1·7	·76	1·7	·76	1·7	·77
·66	1·5	·75	1·5	·75	1·5	·75	1·5	·75	1·5	·75	1·5	·76	1·6	·76	1·6	·77	1·6	·77	1·7	·78
·68	1·5	·76	1·5	·76	1·5	·76	1·5	·76	1·5	·76	1·5	·77	1·5	·77	1·6	·77	1·6	·78	1·6	·78
0·70	1·4	0·76	1·4	0·76	1·4	0·77	1·4	0·77	1·5	0·77	1·5	0·77	1·5	0·78	1·5	0·78	1·6	0·79	1·6	0·79
·75	1·3	·78	1·3	·78	1·3	·79	1·4	·79	1·4	·79	1·4	·79	1·4	·80	1·4	·80	1·5	·81	1·5	·82
·80	1·3	·80	1·3	·80	1·3	·81	1·3	·81	1·3	·81	1·3	·82	1·3	·82	1·4	·83	1·4	·83	1·4	·84
·85	1·2	·83	1·2	·83	1·2	·83	1·2	·83	1·2	·83	1·3	·84	1·3	·84	1·3	·85	1·3	·86	1·3	·87
·90	1·2	·85	1·2	·85	1·2	·85	1·2	·85	1·2	·86	1·2	·86	1·2	·87	1·2	·88	1·3	·89	1·3	·89
·95	1·1	·87	1·1	·88	1·1	·88	1·1	·88	1·1	·88	1·1	·89	1·2	·90	1·2	·90	1·2	·91	1·2	·92
1·00	1·1	·90	1·1	·90	1·1	·90	1·1	·91	1·1	·91	1·1	·92	1·1	·92	1·1	·93	1·2	·94	1·2	·95
1·05	1·0	·93	1·0	·93	1·1	·93	1·1	·93	1·1	·94	1·1	·95	1·1	·95	1·1	·96	1·1	·97	1·1	·99
1·10	1·0	0·96	1·0	0·96	1·0	0·96	1·0	0·96	1·0	0·97	1·0	0·97	1·0	0·98	1·1	0·99	1·1	1·0	1·1	1·0
1·12	1·0	·97	1·0	·97	1·0	·97	1·0	·98	1·0	·98	1·0	·99	1·0	·99	1·0	1·0	1·1	1·0	1·1	1·0
1·14	0·99	·98	0·99	·98	0·99	·98	1·0	·99	1·0	·99	1·0	1·0	1·0	1·0	1·0	1·0	1·0	1·0	1·1	1·0
1·16	·98	·99	·98	·99	·98	·99	0·98	1·0	0·99	1·0	0·99	1·0	1·0	1·0	1·0	1·0	1·0	1·0	1·0	1·0
1·18	·97	1·0	·97	1·0	·97	1·0	·97	1·0	·98	1·0	·98	1·0	0·99	1·0	1·0	1·0	1·0	1·1	1·0	1·1
1·20	0·96	1·0	0·96	1·0	0·96	1·0	0·96	1·0	0·97	1·0	0·97	1·0	0·98	1·0	0·99	1·1	1·0	1·1	1·0	1·1
1·22	·95	1·0	·95	1·0	·95	1·0	·95	1·0	·96	1·0	·96	1·1	·97	1·1	·98	1·1	0·99	1·1	1·0	1·1
1·24	·94	1·0	·94	1·0	·94	1·0	·94	1·1	·95	1·1	·95	1·1	·96	1·1	·97	1·1	·98	1·1	0·99	1·1
1·26	·93	1·1	·93	1·1	·93	1·1	·93	1·1	·94	1·1	·94	1·1	·95	1·1	·96	1·1	·97	1·1	·98	1·1
1·28	·92	1·1	·92	1·1	·92	1·1	·92	1·1	·93	1·1	·93	1·1	·94	1·1	·95	1·1	·96	1·1	·97	1·1
1·30	0·91	1·1	0·91	1·1	0·91	1·1	0·91	1·1	0·92	1·1	0·92	1·1	0·93	1·1	0·94	1·1	0·95	1·1	0·96	1·2
1·32	·90	1·1	·90	1·1	·90	1·1	·90	1·1	·91	1·1	·91	1·1	·92	1·1	·93	1·1	·94	1·2	·95	1·2
1·34	·89	1·1	·89	1·1	·89	1·1	·90	1·1	·90	1·1	·90	1·1	·91	1·1	·92	1·2	·93	1·2	·94	1·2
1·36	·88	1·1	·88	1·1	·89	1·1	·89	1·1	·89	1·1	·90	1·1	·90	1·2	·91	1·2	·92	1·2	·93	1·2
1·38	·88	1·1	·88	1·1	·88	1·1	·88	1·1	·88	1·2	·89	1·2	·89	1·2	·90	1·2	·91	1·2	·92	1·2
1·40	0·87	1·2	0·87	1·2	0·87	1·2	0·87	1·2	0·88	1·2	0·88	1·2	0·88	1·2	0·89	1·2	0·90	1·2	0·91	1·2
1·42	·86	1·2	·86	1·2	·86	1·2	·86	1·2	·87	1·2	·87	1·2	·88	1·2	·88	1·2	·89	1·2	·90	1·3
1·44	·85	1·2	·85	1·2	·85	1·2	·86	1·2	·86	1·2	·86	1·2	·87	1·2	·87	1·2	·88	1·3	·89	1·3
1·46	·84	1·2	·84	1·2	·85	1·2	·85	1·2	·85	1·2	·86	1·2	·86	1·2	·87	1·3	·87	1·3	·88	1·3
1·48	·83	1·2	·83	1·2	·84	1·2	·84	1·2	·84	1·2	·85	1·2	·85	1·3	·86	1·3	·86	1·3	·87	1·3
1·50	0·83	1·2	0·83	1·2	0·84	1·2	0·84	1·2	0·84	1·2	0·84	1·3	0·85	1·3	0·85	1·3	0·86	1·3	0·87	1·3
1·52	·83	1·2	·83	1·2	·83	1·3	·83	1·3	·83	1·3	·84	1·3	·84	1·3	·85	1·3	·85	1·3	·86	1·3
1·54	·82	1·3	·82	1·3	·82	1·3	·82	1·3	·83	1·3	·83	1·3	·83	1·3	·84	1·3	·84	1·3	·85	1·4
1·56	·81	1·3	·82	1·3	·82	1·3	·82	1·3	·82	1·3	·82	1·3	·83	1·3	·83	1·3	·84	1·4	·85	1·4
1·58	·81	1·3	·81	1·3	·81	1·3	·81	1·3	·81	1·3	·81	1·3	·82	1·3	·82	1·4	·83	1·4	·84	1·4
1·60	0·80	1·3	0·80	1·3	0·80	1·3	0·81	1·3	0·81	1·3	0·81	1·3	0·81	1·4	0·82	1·4	0·82	1·4	0·83	1·4
1·62	·80	1·3	·80	1·3	·80	1·3	·80	1·3	·80	1·3	·80	1·4	·81	1·4	·81	1·4	·82	1·4	·82	1·4
1·64	·79	1·3	·79	1·3	·79	1·4	·79	1·4	·80	1·4	·80	1·4	·80	1·4	·81	1·4	·81	1·4	·82	1·4
1·66	·79	1·4	·79	1·4	·79	1·4	·79	1·4	·79	1·4	·79	1·4	·80	1·4	·80	1·4	·80	1·4	·81	1·5
1·68	·78	1·4	·78	1·4	·78	1·4	·78	1·4	·79	1·4	·79	1·4	·79	1·4	·79	1·4	·80	1·5	·80	1·5
1·70	0·78	1·4	0·78	1·4	0·78	1·4	0·78	1·4	0·78	1·4	0·78	1·4	0·78	1·4	0·79	1·5	0·79	1·5	0·80	1·5
1·72	·77	1·4	·77	1·4	·77	1·4	·77	1·4	·77	1·4	·78	1·4	·78	1·5	·78	1·5	·79	1·5	·79	1·5
1·74	·77	1·4	·77	1·4	·77	1·5	·77	1·4	·77	1·5	·77	1·5	·77	1·5	·78	1·5	·78	1·5	·79	1·5
1·76	·77	1·5	·77	1·5	·77	1·5	·77	1·5	·77	1·5	·77	1·5	·77	1·5	·78	1·5	·78	1·5	·78	1·6
1·78	·76	1·5	·76	1·5	·76	1·5	·76	1·5	·76	1·5	·76	1·5	·76	1·5	·77	1·5	·77	1·6	·77	1·6
1·80	0·75	1·5	0·75	1·5	0·75	1·5	0·75	1·5	0·76	1·5	0·76	1·5	0·76	1·5	0·76	1·5	0·77	1·6	0·77	1·6
1·82	·75	1·5	·75	1·5	·75	1·5	·75	1·5	·75	1·5	·75	1·5	·76	1·5	·76	1·6	·76	1·6	·77	1·6
1·84	·74	1·5	·74	1·5	·74	1·5	·74	1·5	·75	1·5	·75	1·6	·75	1·6	·75	1·6	·76	1·6	·76	1·7
1·86	·74	1·5	·74	1·5	·74	1·6	·74	1·6	·74	1·6	·75	1·6	·75	1·6	·75	1·6	·75	1·6	·76	1·7
1·88	·73	1·6	·73	1·6	·73	1·6	·73	1·6	·73	1·6	·74	1·6	·74	1·6	·74	1·6	·75	1·6	·75	1·7
1·90	0·73	1·6	0·73	1·6	0·73	1·6	0·73	1·6	0·73	1·6	0·73	1·6	0·74	1·6	0·74	1·6	0·74	1·7	0·74	1·7
1·95	·72	1·6	·72	1·6	·72	1·6	·72	1·6	·72	1·6	·72	1·7	·73	1·7	·73	1·7	·73	1·7	·73	1·7
2·00	·71	1·7	·71	1·7	·71	1·7	·71	1·7	·71	1·7	·71	1·7	·72	1·7	·72	1·7	·72	1·8	·72	1·8
2·05	·70	1·7	·70	1·7	·70	1·7	·70	1·7	·70	1·7	·70	1·8	·71	1·8	·71	1·8	·71	1·8	·71	1·8
2·10	·69	1·8	·69	1·8	·69	1·8	·69	1·8	·70	1·8	·70	1·8	·70	1·8	·70	1·8	·70	1·9	·70	1·9
2·15	·68	1·8	·68	1·8	·68	1·8	·69	1·8	·69	1·9	·69	1·9	·69	1·9	·69	1·9	·69	1·9	·69	2·0
2·20	0·67	1·9	0·68	1·9	0·68	1·9	0·68	1·9	0·68	1·9	0·68	1·9	0·68	1·9	0·68	2·0	0·68	2·0	0·68	2·0
2·25	·67	2·0	·67	2·0	·67	2·0	·67	2·0	·67	2·0	·67	2·0	·67	2·0	·67	2·0	·67	2·0	·67	2·1
2·30	·66	2·0	·66	2·0	·66	2·0	·66	2·0	·66	2·0	·66	2·0	·66	2·0	·66	2·1	·66	2·1	·67	2·1
2·35	·65	2·1	·65	2·1	·65	2·1	·65	2·1	·65	2·1	·66	2·1	·66	2·1	·66	2·1	·66	2·2	·66	2·2
2·40	·64	2·1	·64	2·1	·65	2·1	·65	2·1	·65	2·1	·65	2·1	·65	2·1	·65	2·2	·65	2·2	·65	2·3
2·45	·64	2·2	·64	2·2	·64	2·2	·64	2·2	·64	2·2	·64	2·2	·64	2·2	·64	2·2	·64	2·3	·64	2·3
2·50	·63	2·3	·63	2·3	·63	2·3	·64	2·2	·64	2·2	·64	2·2	·64	2·3	·64	2·3	·63	2·3	·63	2·4
$\gamma \rightarrow$	2·0		2·1		2·2		2·3		2·4		2·5		2·6		2·7		2·8		2·9	

Five doses (cont.). Standard error of estimates, $s\sqrt{n}$

	γ																	
	3·0		3·1		3·2		3·25		3·3		3·35		3·4		3·45		3·5	
β	$s(\hat{\gamma})$	$s(\hat{\beta})$	$s(\hat{\gamma})$	$s(\hat{\beta})$	$s(\hat{\gamma})$	$s(\hat{\beta})$	$s(\hat{\gamma})$	$s(\hat{\beta})$	$s(\hat{\gamma})$	$s(\hat{\beta})$	$s(\hat{\gamma})$	$s(\hat{\beta})$	$s(\hat{\gamma})$	$s(\hat{\beta})$	$s(\hat{\gamma})$	$w(\hat{\beta})$	$s(\hat{\gamma})$	$s(\hat{\beta})$
0·40	2·8	0·69	2·9	0·69	3·0	0·69	3·0	0·69	3·1	0·70	3·1	0·70	3·2	0·70	3·2	0·70	3·3	0·70
·42	2·6	·69	2·7	·70	2·8	·70	2·9	·70	2·9	·70	3·0	·70	3·0	·71	3·1	·71	3·1	·71
·44	2·5	·70	2·6	·70	2·7	·71	2·8	·71	2·8	·71	2·9	·71	2·9	·71	3·0	·72	3·0	·72
·46	2·4	·71	2·5	·71	2·6	·71	2·6	·71	2·7	·72	2·7	·72	2·8	·72	2·8	·72	2·9	·73
·48	2·3	·71	2·4	·72	2·5	·72	2·5	·72	2·6	·72	2·6	·73	2·7	·73	2·7	·73	2·8	·73
0·50	2·2	0·72	2·3	0·72	2·4	0·73	2·4	0·73	2·5	0·73	2·5	0·73	2·6	0·74	2·6	0·74	2·7	0·74
·52	2·2	·73	2·2	·73	2·3	·73	2·3	·74	2·4	·74	2·4	·74	2·5	·75	2·5	·75	2·6	·75
·54	2·1	·73	2·2	·74	2·2	·74	2·3	·75	2·3	·75	2·3	·75	2·4	·75	2·4	·76	2·5	·76
·56	2·0	·74	2·1	·75	2·1	·75	2·2	·75	2·2	·76	2·3	·76	2·3	·76	2·3	·77	2·4	·77
·58	2·0	·75	2·0	·75	2·1	·76	2·1	·76	2·1	·77	2·2	·77	2·2	·77	2·3	·78	2·3	·78
0·60	1·9	0·76	1·9	0·76	2·0	0·77	2·0	0·77	2·1	0·78	2·1	0·78	2·2	0·78	2·2	0·79	2·2	0·79
·62	1·8	·76	1·9	·77	1·9	·78	2·0	·78	2·0	·78	2·0	·79	2·1	·79	2·1	·80	2·2	·80
·64	1·8	·77	1·8	·78	1·9	·79	1·9	·79	2·0	·79	2·0	·80	2·0	·80	2·1	·81	2·1	·81
·66	1·7	·78	1·8	·79	1·8	·80	1·9	·80	1·9	·80	1·9	·81	2·0	·81	2·0	·82	2·0	·82
·68	1·7	·79	1·7	·80	1·8	·81	1·8	·81	1·8	·81	1·9	·82	1·9	·82	1·9	·83	2·0	·83
0·70	1·6	0·80	1·7	0·81	1·7	0·82	1·8	0·82	1·8	0·83	1·8	0·83	1·9	0·84	1·9	0·84	1·9	0·85
·75	1·5	·82	1·6	·83	1·6	·84	1·7	·85	1·7	·85	1·7	·86	1·7	·86	1·8	·87	1·8	·88
·80	1·5	·85	1·5	·86	1·5	·87	1·6	·88	1·6	·88	1·6	·89	1·6	·90	1·7	·90	1·7	·91
·85	1·4	·88	1·4	·89	1·4	·90	1·5	·91	1·5	·91	1·5	·92	1·6	·93	1·6	·94	1·6	·94
·90	1·3	·91	1·3	·92	1·4	·93	1·4	·94	1·4	·95	1·5	·95	1·5	·96	1·5	·97	1·5	·98
·95	1·3	·94	1·3	·95	1·3	·96	1·3	·97	1·4	·98	1·4	·99	1·4	1·0	1·4	1·0	1·5	1·0
1·00	1·2	·97	1·2	·98	1·3	1·0	1·3	1·0	1·3	1·0	1·3	1·0	1·3	1·0	1·4	1·1	1·4	1·1
1·05	1·1	1·0	1·2	1·0	1·2	1·0	1·2	1·0	1·2	1·1	1·3	1·1	1·3	1·1	1·3	1·1	1·3	1·1
1·10	1·1	1·0	1·1	1·1	1·2	1·1	1·2	1·1	1·2	1·1	1·2	1·1	1·2	1·1	1·3	1·1	1·3	1·1
1·12	1·1	1·0	1·1	1·1	1·1	1·1	1·2	1·1	1·2	1·1	1·2	1·1	1·2	1·1	1·2	1·1	1·3	1·2
1·14	1·1	1·1	1·1	1·1	1·1	1·1	1·1	1·1	1·2	1·1	1·2	1·1	1·2	1·1	1·2	1·2	1·2	1·2
1·16	1·0	1·1	1·1	1·1	1·1	1·1	1·1	1·1	1·1	1·1	1·2	1·2	1·2	1·2	1·2	1·2	1·2	1·2
1·18	1·0	1·1	1·1	1·1	1·1	1·1	1·1	1·1	1·1	1·2	1·1	1·2	1·2	1·2	1·2	1·2	1·2	1·2
1·20	1·0	1·1	1·1	1·1	1·1	1·1	1·1	1·2	1·1	1·2	1·1	1·2	1·1	1·2	1·2	1·2	1·2	1·2
1·22	1·0	1·1	1·0	1·1	1·1	1·2	1·1	1·2	1·1	1·2	1·1	1·2	1·1	1·2	1·1	1·2	1·2	1·2
1·24	1·0	1·1	1·0	1·2	1·1	1·2	1·1	1·2	1·1	1·2	1·1	1·2	1·1	1·2	1·1	1·3	1·1	1·3
1·26	0·99	1·2	1·0	1·2	1·0	1·2	1·0	1·2	1·1	1·2	1·1	1·2	1·1	1·3	1·1	1·3	1·1	1·3
1·28	·98	1·2	1·0	1·2	1·0	1·2	1·0	1·2	1·0	1·2	1·1	1·3	1·1	1·3	1·1	1·3	1·1	1·3
1·30	0·97	1·2	0·99	1·2	1·0	1·2	1·0	1·2	1·0	1·3	1·1	1·3	1·1	1·3	1·1	1·3	1·1	1·3
1·32	·96	1·2	·98	1·2	0·99	1·2	1·0	1·3	1·0	1·3	1·0	1·3	1·0	1·3	1·1	1·3	1·1	1·3
1·34	·95	1·2	·97	1·2	·99	1·3	1·0	1·3	1·0	1·3	1·0	1·3	1·0	1·3	1·1	1·3	1·1	1·4
1·36	·94	1·2	·96	1·3	·98	1·3	0·99	1·3	1·0	1·3	1·0	1·3	1·0	1·3	1·0	1·4	1·1	1·4
1·38	·93	1·2	·95	1·3	·96	1·3	·98	1·3	0·99	1·3	1·0	1·3	1·0	1·4	1·0	1·4	1·0	1·4
1·40	0·92	1·3	0·94	1·3	0·95	1·3	0·96	1·3	0·98	1·3	0·99	1·4	1·0	1·4	1·0	1·4	1·0	1·4
1·42	·91	1·3	·93	1·3	·94	1·3	·95	1·3	·96	1·4	·98	1·4	0·99	1·4	1·0	1·4	1·0	1·4
1·44	·90	1·3	·92	1·3	·93	1·3	·94	1·4	·95	1·4	·97	1·4	·98	1·4	0·99	1·4	1·0	1·5
1·46	·90	1·3	·91	1·3	·92	1·4	·93	1·4	·94	1·4	·95	1·4	·97	1·4	·98	1·5	0·99	1·5
1·48	·89	1·3	·90	1·4	·91	1·4	·92	1·4	·93	1·4	·94	1·4	·96	1·5	·97	1·5	·98	1·5
1·50	0·88	1·3	0·89	1·4	0·90	1·4	0·91	1·4	0·92	1·4	0·93	1·5	0·95	1·5	0·96	1·5	0·97	1·5
1·52	·87	1·4	·88	1·4	·90	1·4	·90	1·4	·91	1·5	·92	1·5	·94	1·5	·95	1·5	·96	1·5
1·54	·86	1·4	·87	1·4	·89	1·4	·90	1·5	·90	1·5	·92	1·5	·93	1·5	·94	1·5	·95	1·6
1·56	·86	1·4	·87	1·4	·88	1·5	·89	1·5	·90	1·5	·91	1·5	·92	1·5	·93	1·6	·94	1·6
1·58	·85	1·4	·86	1·4	·87	1·5	·88	1·5	·89	1·5	·90	1·5	·91	1·6	·92	1·6	·93	1·6
1·60	0·84	1·4	0·85	1·5	0·86	1·5	0·87	1·5	0·88	1·5	0·89	1·6	0·90	1·6	0·91	1·6	0·92	1·6
1·62	·83	1·5	·84	1·5	·85	1·5	·86	1·5	·87	1·6	·88	1·6	·89	1·6	·90	1·6	·91	1·6
1·64	·82	1·5	·83	1·5	·85	1·5	·85	1·6	·86	1·6	·87	1·6	·88	1·6	·89	1·6	·91	1·7
1·66	·82	1·5	·83	1·5	·84	1·6	·84	1·6	·85	1·6	·86	1·6	·87	1·6	·88	1·7	·90	1·7
1·68	·81	1·5	·82	1·5	·83	1·6	·84	1·6	·85	1·6	·86	1·6	·87	1·7	·88	1·7	·89	1·7
1·70	0·81	1·5	0·81	1·6	0·83	1·6	0·83	1·6	0·84	1·6	0·85	1·7	0·86	1·7	0·87	1·7	0·88	1·7
1·72	·80	1·5	·81	1·6	·82	1·6	·82	1·6	·83	1·7	·84	1·7	·85	1·7	·86	1·7	·87	1·8
1·74	·79	1·6	·80	1·6	·81	1·6	·82	1·7	·83	1·7	·83	1·7	·84	1·7	·85	1·7	·86	1·8
1·76	·79	1·6	·80	1·6	·80	1·7	·81	1·7	·82	1·7	·82	1·7	·83	1·7	·84	1·8	·85	1·8
1·78	·78	1·6	·79	1·6	·79	1·7	·80	1·7	·81	1·7	·81	1·7	·82	1·8	·83	1·8	·84	1·8
1·80	0·78	1·6	0·78	1·7	0·79	1·7	0·80	1·7	0·81	1·7	0·81	1·8	0·82	1·8	0·83	1·8	0·84	1·8
1·82	·77	1·6	·78	1·7	·79	1·7	·79	1·7	·80	1·8	·80	1·8	·81	1·8	·82	1·8	·83	1·9
1·84	·76	1·7	·77	1·7	·78	1·7	·79	1·8	·79	1·8	·80	1·8	·81	1·8	·82	1·9	·83	1·9
1·86	·76	1·7	·76	1·7	·78	1·8	·78	1·8	·78	1·8	·79	1·8	·80	1·8	·81	1·9	·82	1·9
1·88	·75	1·7	·76	1·7	·77	1·8	·77	1·8	·78	1·8	·79	1·8	·80	1·9	·80	1·9	·81	1·9
1·90	0·75	1·7	0·76	1·8	0·76	1·8	0·77	1·8	0·78	1·8	0·78	1·9	0·79	1·9	0·80	1·9	0·81	1·9
1·95	·74	1·8	·74	1·8	·75	1·8	·75	1·9	·76	1·9	·77	1·9	·77	1·9	·78	2·0	·79	2·0
2·00	·73	1·8	·73	1·9	·74	1·9	·74	1·9	·75	1·9	·75	2·0	·76	2·0	·77	2·0	·78	2·0
2·05	·72	1·9	·72	1·9	·73	2·0	·73	2·0	·73	2·0	·74	2·0	·75	2·1	·75	2·1	·76	2·1
2·10	·70	1·9	·71	2·0	·72	2·0	·72	2·0	·72	2·1	·73	2·1	·74	2·1	·74	2·1	·75	2·2
2·15	·70	2·0	·70	2·0	·71	2·1	·71	2·1	·71	2·1	·72	2·1	·72	2·2	·73	2·2	·74	2·2
2·20	0·69	2·0	0·69	2·1	0·70	2·1	0·70	2·1	0·70	2·2	0·71	2·2	0·71	2·2	0·72	2·2	0·72	2·3
2·25	·68	2·1	·68	2·1	·69	2·2	·69	2·2	·69	2·2	·70	2·2	·70	2·3	·71	2·3	·71	2·3
2·30	·67	2·2	·67	2·2	·68	2·2	·68	2·3	·68	2·3	·69	2·3	·69	2·3	·70	2·4	·70	2·4
2·35	·66	2·2	·66	2·3	·67	2·3	·67	2·3	·67	2·3	·68	2·4	·68	2·4	·69	2·4	·69	2·4
2·40	·65	2·3	·65	2·3	·66	2·4	·66	2·4	·67	2·4	·67	2·4	·67	2·4	·68	2·5	·68	2·5
2·45	·64	2·4	·65	2·4	·65	2·4	·66	2·4	·66	2·4	·66	2·5	·67	2·5	·67	2·5	·68	2·6
2·50	·64	2·4	·64	2·5	·64	2·5	·65	2·5	·65	2·5	·65	2·5	·66	2·5	·66	2·6	·67	2·6
$\gamma\rightarrow$	3·0		3·1		3·2		3·25		3·3		3·35		3·4		3·45		3·5	

Six doses. Standard error of estimates, $s\sqrt{n}$

	γ																							
	2·5		2·6		2·7		2·8		2·9		3·0		3·1		3·2		3·3		3·4		3·5		3·6	
β	$s(\frac{x}{y})$	$s(\beta)$	$s(\frac{x}{y})$	$s(\beta)$	$s(\frac{x}{y})$	$s(\beta)$	$s(\frac{x}{y})$	$s(\beta)$	$s(\frac{x}{y})$	$s(\beta)$	$s(\frac{x}{y})$	$s(\beta)$	$s(\frac{x}{y})$	$s(\beta)$	$s(\frac{x}{y})$	$s(\beta)$	$s(\frac{x}{y})$	$s(\beta)$	$s(\frac{x}{y})$	$s(\beta)$	$s(\frac{x}{y})$	$s(\beta)$	$s(\frac{x}{y})$	$s(\beta)$
0·40	2·2	0·53	2·2	0·53	2·2	0·53	2·2	0·53	2·2	0·53	2·2	0·53	2·3	0·53	2·3	0·53	2·3	0·53	2·4	0·53	2·4	0·54	2·5	0·54
·42	2·1	·53	2·1	·53	2·1	·53	2·1	·53	2·1	·53	2·1	·53	2·2	·53	2·2	·54	2·2	·54	2·3	·54	2·3	·54	2·4	·54
·44	2·0	·54	2·0	·54	2·0	·54	2·0	·54	2·0	·54	2·0	·54	2·1	·54	2·1	·54	2·1	·54	2·2	·55	2·2	·55	2·3	·55
·46	1·9	·54	1·9	·54	1·9	·54	1·9	·54	1·9	·54	2·0	·54	2·0	·55	2·0	·55	2·0	·55	2·1	·55	2·1	·55	2·2	·56
·48	1·8	·55	1·8	·55	1·9	·55	1·9	·55	1·9	·55	1·9	·55	1·9	·55	1·9	·55	2·0	·55	2·0	·56	2·1	·56	2·1	·56
0·50	1·8	0·55	1·8	0·55	1·8	0·55	1·8	0·55	1·8	0·56	1·8	0·56	1·8	0·56	1·9	0·56	1·9	0·56	1·9	0·56	2·0	0·57	2·0	0·57
·52	1·7	·56	1·7	·56	1·7	·56	1·7	·56	1·7	·56	1·8	·56	1·8	·56	1·8	·57	1·8	·57	1·9	·57	1·9	·57	1·9	·58
·54	1·7	·57	1·7	·57	1·7	·57	1·7	·57	1·7	·57	1·7	·57	1·7	·57	1·7	·57	1·8	·58	1·8	·58	1·8	·58	1·9	·59
·56	1·6	·57	1·6	·57	1·6	·57	1·6	·57	1·6	·57	1·7	·58	1·7	·58	1·7	·59	1·7	·59	1·7	·59	1·8	·59	1·8	·59
·58	1·6	·58	1·6	·58	1·6	·58	1·6	·58	1·6	·58	1·6	·58	1·6	·59	1·6	·59	1·7	·59	1·7	·59	1·7	·60	1·8	·60
0·60	1·5	0·59	1·5	0·59	1·5	0·59	1·5	0·59	1·5	0·59	1·6	0·59	1·6	0·59	1·6	0·60	1·6	0·60	1·6	0·60	1·7	0·61	1·7	0·61
·62	1·5	·59	1·5	·59	1·5	·59	1·5	·60	1·5	·60	1·5	·60	1·5	·60	1·5	·60	1·6	·61	1·6	·61	1·6	·61	1·7	·62
·64	1·4	·60	1·4	·60	1·5	·60	1·5	·60	1·5	·60	1·5	·61	1·5	·61	1·5	·61	1·5	·61	1·5	·62	1·6	·62	1·6	·63
·66	1·4	·61	1·4	·61	1·4	·61	1·4	·61	1·4	·61	1·4	·61	1·5	·62	1·5	·62	1·5	·62	1·5	·63	1·5	·63	1·6	·64
·68	1·4	·62	1·4	·62	1·4	·62	1·4	·62	1·4	·62	1·4	·62	1·4	·63	1·4	·63	1·5	·63	1·5	·64	1·5	·64	1·5	·65
0·70	1·3	0·62	1·4	0·63	1·4	0·63	1·4	0·63	1·4	0·63	1·4	0·63	1·4	0·63	1·4	0·64	1·4	0·64	1·4	0·65	1·5	0·65	1·5	0·66
·75	1·3	·65	1·3	·65	1·3	·65	1·3	·65	1·3	·65	1·3	·65	1·3	·66	1·3	·66	1·3	·66	1·4	·67	1·4	·68	1·4	·68
·80	1·2	·67	1·2	·67	1·2	·67	1·2	·67	1·2	·67	1·2	·68	1·3	·68	1·2	·68	1·3	·69	1·3	·69	1·3	·70	1·3	·71
·85	1·2	·69	1·2	·69	1·2	·69	1·2	·70	1·2	·70	1·2	·70	1·2	·71	1·2	·71	1·2	·72	·12	·72	1·3	·73	1·3	·74
·90	1·1	·72	1·1	·72	1·1	·72	1·1	·72	1·1	·73	1·1	·73	1·1	·73	1·2	·74	1·2	·74	1·2	·75	1·2	·76	1·2	·77
·95	1·1	·75	1·1	·75	1·1	·75	1·1	·75	1·1	·75	1·1	·76	1·1	·76	1·1	·77	1·1	·77	1·1	·78	1·1	·79	1·2	·80
1·00	1·0	·77	1·0	·78	1·1	·78	1·1	·78	1·1	·78	1·1	·79	1·1	·79	1·1	·80	1·1	·80	1·1	·81	1·1	·82	1·1	·83
1·05	1·0	·80	1·0	·81	1·0	·81	1·0	·81	1·0	·81	1·0	·82	1·0	·82	1·0	·83	1·0	·83	1·1	·84	1·1	·85	1·1	·86
1·10	0·99	0·84	0·99	0·84	0·99	0·84	0·99	0·84	0·99	0·84	1·0	0·85	1·0	0·85	1·0	0·86	1·0	0·87	1·0	0·88	1·0	0·89	1·0	0·90
1·12	·98	·85	·98	·85	·98	·85	·98	·85	·98	·85	0·98	·86	0·99	·87	0·99	·87	1·0	·88	1·0	·89	1·0	·90	1·0	·91
1·14	·97	·86	·97	·86	·97	·86	·97	·87	·97	·87	·97	·87	·98	·88	·98	·89	0·99	·90	1·0	·91	1·0	·92	1·0	·93
1·16	·96	·88	·96	·88	·96	·88	·96	·88	·96	·89	·96	·89	·97	·89	·97	·90	·98	·91	0·98	·92	0·99	·93	1·0	·94
1·18	·95	·89	·95	·89	·95	·89	·95	·89	·95	·90	·95	·90	·96	·91	·96	·92	·97	·92	·97	·93	·98	·94	0·99	·96
1·20	0·94	0·90	0·94	0·91	0·94	0·91	0·94	0·91	0·94	0·91	0·94	0·92	0·95	0·92	0·95	0·93	0·96	0·94	0·96	0·95	0·97	0·96	0·98	0·97
1·22	·93	·92	·93	·92	·93	·92	·93	·92	·93	·93	·93	·94	·94	·94	·94	·94	·95	·95	·95	·96	·96	·97	·97	·99
1·24	·92	·93	·92	·93	·92	·93	·92	·93	·92	·94	·92	·94	·93	·95	·93	·96	·94	·97	·94	·98	·95	·99	·96	1·0
1·26	·91	·95	·91	·95	·91	·95	·91	·95	·91	·95	·91	·96	·92	·97	·92	·97	·93	·98	·93	·99	·94	1·0	·95	1·0
1·28	·90	·96	·90	·96	·90	·96	·90	·97	·90	·97	·91	·97	·91	·98	·91	·99	·92	1·0	·92	1·0	·93	1·0	·94	1·0
1·30	0·89	0·97	0·89	0·98	0·89	0·98	0·90	0·98	0·90	0·98	0·90	0·99	0·90	1·0	0·91	1·0	0·91	1·0	0·91	1·0	0·92	1·0	0·93	1·0
1·32	·88	·99	·88	·99	·89	·99	·89	·99	·89	·99	·89	1·0	·89	1·0	·90	1·0	·90	1·0	·91	1·0	·91	1·1	·92	1·1
1·34	·88	1·0	·88	1·0	·88	1·0	·88	1·0	·88	1·0	·88	1·0	·88	1·0	·89	1·0	·89	1·1	·90	1·1	·90	1·1	·91	1·1
1·36	·87	1·0	·87	1·0	·87	1·0	·87	1·0	·87	1·0	·87	1·0	·88	1·0	·88	1·1	·88	1·1	·89	1·1	·89	1·1	·90	1·1
1·38	·86	1·0	·86	1·0	·86	1·0	·86	1·0	·86	1·0	·87	1·1	·87	1·1	·87	1·1	·88	1·1	·88	1·1	·89	1·1	·89	1·1
1·40	0·86	1·1	0·86	1·1	0·86	1·1	0·86	1·1	0·86	1·1	0·86	1·1	0·86	1·1	0·87	1·1	0·87	1·1	0·87	1·1	0·88	1·1	0·88	1·1
1·42	·85	1·1	·85	1·1	·85	1·1	·85	1·1	·85	1·1	·85	1·1	·86	1·1	·86	1·1	·86	1·1	·86	1·1	·87	1·1	·87	1·1
1·44	·84	1·1	·84	1·1	·84	1·1	·84	1·1	·85	1·1	·85	1·1	·85	1·1	·85	1·1	·85	1·1	·86	1·1	·86	1·1	·87	1·2
1·46	·84	1·1	·84	1·1	·84	1·1	·84	1·1	·84	1·1	·84	1·1	·84	1·1	·84	1·1	·85	1·1	·85	1·2	·85	1·2	·86	1·2
1·48	·83	1·1	·83	1·1	·83	1·1	·83	1·1	·83	1·1	·83	1·1	·83	1·1	·84	1·1	·84	1·2	·84	1·2	·85	1·2	·85	1·2
1·50	0·82	1·1	0·83	1·1	0·83	1·1	0·83	1·1	0·83	1·1	0·83	1·2	0·83	1·2	0·83	1·2	0·83	1·2	0·84	1·2	0·84	1·2	0·85	1·2
1·52	·82	1·2	·82	1·2	·82	1·2	·82	1·2	·82	1·2	·82	1·2	·82	1·2	·83	1·2	·83	1·2	·83	1·2	·83	1·2	·84	1·2
1·54	·81	1·2	·81	1·2	·81	1·2	·81	1·2	·82	1·2	·82	1·2	·82	1·2	·82	1·2	·82	1·2	·82	1·2	·83	1·2	·83	1·3
1·56	·81	1·2	·81	1·2	·81	1·2	·81	1·2	·81	1·2	·81	1·2	·81	1·2	·81	1·2	·82	1·2	·82	1·2	·82	1·3	·83	1·3
1·58	·80	1·2	·80	1·2	·80	1·2	·80	1·2	·80	1·2	·80	1·2	·81	1·2	·81	1·2	·81	1·2	·81	1·3	·81	1·3	·82	1·3
1·60	0·80	1·2	0·80	1·2	0·80	1·2	0·80	1·2	0·80	1·2	0·80	1·2	0·80	1·2	0·80	1·3	0·80	1·3	0·81	1·3	0·81	1·3	0·81	1·3
1·62	·79	1·2	·79	1·2	·79	1·2	·79	1·3	·80	1·3	·80	1·3	·80	1·3	·80	1·3	·80	1·3	·80	1·3	·80	1·3	·80	1·3
1·64	·78	1·3	·79	1·3	·79	1·3	·79	1·3	·79	1·3	·79	1·3	·79	1·3	·79	1·3	·79	1·3	·80	1·3	·80	1·3	·80	1·3
1·66	·78	1·3	·78	1·3	·78	1·3	·78	1·3	·78	1·3	·79	1·3	·79	1·3	·79	1·3	·79	1·3	·79	1·3	·79	1·3	·80	1·4
1·68	·78	1·3	·78	1·3	·78	1·3	·78	1·3	·78	1·3	·78	1·3	·78	1·3	·78	1·3	·78	1·3	·78	1·4	·79	1·4	·79	1·4
1·70	0·77	1·3	0·77	1·3	0·77	1·3	0·77	1·3	0·77	1·3	0·77	1·3	0·77	1·3	0·78	1·3	0·78	1·4	0·78	1·4	0·78	1·4	0·78	1·4
1·72	·77	1·3	·77	1·3	·77	1·3	·77	1·3	·77	1·3	·77	1·3	·77	1·4	·77	1·4	·77	1·4	·78	1·4	·78	1·4	·78	1·4
1·74	·76	1·4	·76	1·4	·76	1·4	·77	1·4	·77	1·4	·77	1·4	·77	1·4	·77	1·4	·77	1·4	·77	1·4	·77	1·4	·78	1·4
1·76	·76	1·4	·76	1·4	·76	1·4	·76	1·4	·76	1·4	·76	1·4	·76	1·4	·76	1·4	·77	1·4	·77	1·4	·77	1·4	·77	1·5
1·78	·75	1·4	·75	1·4	·75	1·4	·75	1·4	·75	1·4	·75	1·4	·75	1·4	·75	1·4	·76	1·4	·76	1·4	·76	1·5	·76	1·5
1·80	0·75	1·4	0·75	1·4	0·75	1·4	0·75	1·4	0·75	1·4	0·75	1·4	0·75	1·4	0·75	1·4	0·75	1·4	0·75	1·4	0·76	1·5	0·76	1·5
1·82	·74	1·4	·74	1·4	·74	1·4	·74	1·4	·74	1·4	·74	1·4	·74	1·5	·74	1·5	·75	1·5	·75	1·5	·75	1·5	·75	1·5
1·84	·74	1·4	·74	1·4	·74	1·5	·74	1·5	·74	1·5	·74	1·5	·74	1·5	·74	1·5	·74	1·5	·75	1·5	·75	1·5	·75	1·5
1·86	·74	1·5	·74	1·5	·74	1·5	·74	1·5	·74	1·5	·74	1·5	·74	1·5	·74	1·5	·74	1·5	·74	1·5	·74	1·6	·74	1·6
1·88	·73	1·5	·73	1·5	·73	1·5	·73	1·5	·73	1·5	·73	1·5	·73	1·5	·73	1·5	·74	1·5	·74	1·5	·74	1·6	·74	1·6
1·90	0·73	1·5	0·73	1·5	0·73	1·5	0·73	1·5	0·73	1·5	0·73	1·5	0·73	1·5	0·73	1·5	0·73	1·6	0·73	1·6	0·73	1·6	0·74	1·6
1·92	·72	1·5	·72	1·5	·73	1·5	·73	1·5	·73	1·5	·73	1·6	·73	1·6	·73	1·6	·73	1·6	·73	1·6	·73	1·6	·73	1·6
1·94	·72	1·5	·72	1·6	·72	1·6	·72	1·6	·72	1·6	·72	1·6	·72	1·6	·72	1·6	·72	1·6	·72	1·6	·73	1·6	·73	1·6
1·96	·72	1·6	·72	1·6	·72	1·6	·72	1·6	·72	1·6	·72	1·6	·72	1·6	·72	1·6	·72	1·6	·72	1·6	·73	1·6	·73	1·7
1·98	·71	1·6	·71	1·6	·71	1·6	·71	1·6	·71	1·6	·71	1·6	·71	1·6	·71	1·6	·72	1·6	·72	1·6	·72	1·7	·72	1·7
2·00	0·71	1·6	0·71	1·6	0·71	1·6	0·71	1·6	0·71	1·6	0·71	1·6	0·71	1·6	0·71	1·6	0·71	1·7	0·71	1·7	0·71	1·7	0·72	1·7
$\gamma\rightarrow$	2·5		2·6		2·7		2·8		2·9		3·0		3·1		3·2		3·3		3·4		3·5		3·6	

Six doses (cont.). Standard error of estimates, $s\sqrt{n}$

	γ																					
	3·7		3·8		3·9		4·0		4·1		4·2		4·25		4·3		4·35		4·4		4·45		4·5	
β	$s(\frac{\alpha}{\gamma})$	$s(\beta)$	$s(\frac{\alpha}{\gamma})$	$s(\beta)$	$s(\frac{\alpha}{\gamma})$	$s(\beta)$	$s(\frac{\alpha}{\gamma})$	$s(\beta)$	$s(\frac{\alpha}{\gamma})$	$s(\beta)$	$s(\frac{\alpha}{\gamma})$	$s(\beta)$	$s(\frac{\alpha}{\gamma})$	$s(\beta)$	$s(\frac{\alpha}{\gamma})$	$s(\beta)$	$s(\frac{\alpha}{\gamma})$	$s(\beta)$	$s(\frac{\alpha}{\gamma})$	$s(\beta)$	$s(\frac{\alpha}{\gamma})$	$s(\beta)$
0·40	2·6	0·54	2·6	0·54	2·7	0·54	2·8	0·55	2·9	0·55	2·9	0·55	3·0	0·55	3·0	0·56	3·1	0·56	3·1	0·56	3·2	0·56	3·2	0·56
·42	2·4	·55	2·5	·55	2·6	·55	2·6	·55	2·7	·56	2·8	·56	2·8	·56	2·9	·56	2·9	·57	3·0	·57	3·0	·57	3·1	·57
·44	2·3	·55	2·4	·56	2·5	·56	2·5	·56	2·6	·56	2·7	·57	2·7	·57	2·8	·57	2·8	·57	2·8	·58	2·9	·58	2·9	·58
·46	2·2	·56	2·3	·56	2·4	·57	2·4	·57	2·5	·57	2·6	·58	2·6	·58	2·6	·58	2·7	·58	2·7	·59	2·8	·59	2·8	·59
·48	2·1	·57	2·2	·57	2·3	·57	2·3	·58	2·4	·59	2·5	·59	2·5	·59	2·5	·59	2·6	·59	2·6	·60	2·7	·60	2·7	·60
0·50	2·1	0·57	2·1	0·58	2·2	0·58	2·2	0·59	2·3	0·59	2·4	0·59	2·4	0·60	2·4	0·60	2·5	0·60	2·5	0·60	2·6	0·61	2·6	0·61
·52	2·0	·58	2·0	·58	2·1	·59	2·2	·59	2·2	·60	2·3	·60	2·3	·61	2·3	·61	2·4	·61	2·4	·61	2·5	·62	2·5	·62
·54	1·9	·59	2·0	·59	2·1	·60	2·1	·60	2·1	·61	2·2	·61	2·2	·62	2·3	·62	2·3	·62	2·3	·63	2·4	·63	2·4	·63
·56	1·9	·60	1·9	·60	2·0	·61	2·0	·61	2·1	·62	2·1	·62	2·2	·63	2·2	·63	2·2	·63	2·3	·64	2·3	·64	2·3	·64
·58	1·8	·61	1·8	·61	1·9	·62	1·9	·62	2·0	·63	2·1	·63	2·1	·64	2·1	·64	2·2	·64	2·2	·65	2·2	·65	2·3	·66
0·60	1·7	0·61	1·8	0·62	1·8	0·62	1·9	0·63	1·9	0·64	2·0	0·64	2·0	0·65	2·1	0·65	2·1	0·66	2·1	0·66	2·2	0·66	2·2	0·67
·62	1·7	·62	1·7	·63	1·8	·63	1·8	·64	1·9	·65	1·9	·65	2·0	·66	2·0	·66	2·0	·67	2·1	·67	2·1	·68	2·1	·68
·64	1·6	·63	1·7	·64	1·7	·64	1·8	·65	1·8	·66	1·9	·67	1·9	·67	1·9	·67	2·0	·68	2·0	·68	2·0	·69	2·1	·69
·66	1·6	·64	1·6	·65	1·7	·65	1·7	·66	1·8	·67	1·8	·68	1·9	·68	1·9	·69	1·9	·69	1·9	·70	2·0	·70	2·0	·71
·68	1·6	·65	1·6	·66	1·6	·67	1·7	·67	1·7	·68	1·8	·69	1·8	·69	1·8	·70	1·9	·70	1·9	·71	1·9	·71	2·0	·72
0·70	1·5	0·66	1·6	0·67	1·6	0·68	1·6	0·68	1·7	0·69	1·7	0·70	1·8	0·71	1·8	0·71	1·8	0·71	1·8	0·72	1·9	0·72	1·9	0·73
·75	1·4	·69	1·5	·70	1·5	·70	1·5	·71	1·6	·72	1·6	·73	1·7	·74	1·7	·74	1·7	·75	1·7	·76	1·8	·76	1·8	·77
·80	1·4	·72	1·4	·72	1·4	·73	1·5	·74	1·5	·75	1·5	·77	1·6	·77	1·6	·78	1·6	·79	1·6	·79	1·7	·80	1·7	·81
·85	1·3	·74	1·3	·75	1·3	·76	1·4	·77	1·4	·79	1·4	·80	1·5	·81	1·5	·82	1·5	·82	1·5	·83	1·6	·84	1·6	·85
·90	1·2	·78	1·3	·79	1·3	·80	1·3	·81	1·3	·82	1·4	·84	1·4	·84	1·4	·85	1·4	·86	1·5	·87	1·5	·88	1·5	·89
·95	1·2	·81	1·2	·82	1·2	·83	1·3	·84	1·3	·86	1·3	·88	1·3	·88	1·3	·89	1·4	·90	1·4	·91	1·4	·92	1·4	·93
1·00	1·1	·84	1·2	·85	1·2	·87	1·2	·88	1·2	·90	1·3	·91	1·3	·92	1·3	·93	1·3	·94	1·3	·95	1·4	·97	1·4	·98
1·05	1·1	·88	1·1	·89	1·1	·90	1·2	·92	1·2	·94	1·2	·96	1·2	·97	1·2	·98	1·3	·99	1·3	1·0	1·3	1·0	1·3	1·0
1·10	1·1	0·91	1·1	0·92	1·1	0·94	1·1	0·96	1·1	0·98	1·2	1·0	1·2	1·0	1·2	1·0	1·2	1·0	1·2	1·0	1·2	1·0	1·3	1·0
1·12	1·0	·93	1·0	·94	1·1	·96	1·1	·97	1·1	·99	1·1	1·0	1·2	1·0	1·2	1·0	1·2	1·0	1·2	1·1	1·2	1·1	1·2	1·1
1·14	1·0	·94	1·0	·95	1·1	·97	1·1	·99	1·1	1·0	1·1	1·0	1·1	1·0	1·2	1·0	1·2	1·1	1·2	1·1	1·2	1·1	1·2	1·1
1·16	1·0	·96	1·0	·97	1·0	·99	1·1	1·0	1·1	1·0	1·1	1·0	1·1	1·1	1·1	1·1	1·2	1·1	1·2	1·1	1·2	1·1	1·2	1·1
1·18	1·0	·97	1·0	·99	1·0	1·0	1·0	1·0	1·1	1·0	1·1	1·1	1·1	1·1	1·1	1·1	1·1	1·1	1·2	1·1	1·2	1·1	1·2	1·1
1·20	0·99	0·99	1·0	1·0	1·0	1·0	1·0	1·0	1·1	1·1	1·1	1·1	1·1	1·1	1·1	1·1	1·1	1·1	1·1	1·1	1·2	1·2	1·2	1·2
1·22	·98	1·0	0·99	1·0	1·0	1·0	1·1	1·0	1·1	1·1	1·1	1·1	1·1	1·1	1·1	1·1	1·1	1·1	1·1	1·2	1·1	1·2	1·2	1·2
1·24	·97	1·0	·98	1·0	0·99	1·1	1·0	1·1	1·0	1·1	1·0	1·1	1·1	1·1	1·1	1·1	1·1	1·2	1·1	1·2	1·1	1·2	1·1	1·2
1·26	·96	1·0	·97	1·0	·98	1·1	0·99	1·1	1·0	1·1	1·0	1·1	1·0	1·2	1·1	1·2	1·1	1·2	1·1	1·2	1·1	1·2	1·1	1·2
1·28	·95	1·0	·96	1·1	·97	1·1	·98	1·1	1·0	1·1	1·0	1·2	1·0	1·2	1·0	1·2	1·1	1·2	1·1	1·2	1·1	1·2	1·1	1·3
1·30	0·94	1·1	0·95	1·1	0·96	1·1	0·97	1·1	0·99	1·1	1·0	1·2	1·0	1·2	1·0	1·2	1·0	1·2	1·1	1·2	1·1	1·3	1·1	1·3
1·32	·93	1·1	·94	1·1	·95	1·1	·96	1·1	·98	1·2	1·0	1·2	1·0	1·2	1·0	1·2	1·0	1·2	1·1	1·3	1·1	1·3	1·1	1·3
1·34	·92	1·1	·93	1·1	·94	1·1	·95	1·2	·97	1·2	0·99	1·2	1·0	1·2	1·0	1·2	1·0	1·3	1·0	1·3	1·1	1·3	1·1	1·3
1·36	·91	1·1	·92	1·1	·93	1·2	·94	1·2	·96	1·2	·98	1·2	0·98	1·2	1·0	1·3	1·0	1·3	1·0	1·3	1·0	1·3	1·1	1·3
1·38	·90	1·1	·91	1·1	·92	1·2	·93	1·2	·95	1·2	·96	1·2	·97	1·3	0·99	1·3	1·0	1·3	1·0	1·3	1·0	1·3	1·0	1·4
1·40	0·89	1·1	0·90	1·2	0·91	1·2	0·92	1·2	0·94	1·2	0·95	1·3	0·96	1·3	0·98	1·3	0·99	1·3	1·0	1·3	1·0	1·4	1·0	1·4
1·42	·88	1·2	·89	1·2	·90	1·2	·91	1·2	·93	1·3	·94	1·3	·95	1·3	·96	1·3	·97	1·3	0·99	1·4	1·0	1·4	1·0	1·4
1·44	·87	1·2	·88	1·2	·89	1·2	·90	1·2	·92	1·3	·93	1·3	·94	1·3	·95	1·3	·96	1·4	·98	1·4	0·99	1·4	1·0	1·4
1·46	·86	1·2	·87	1·2	·88	1·2	·89	1·3	·91	1·3	·92	1·3	·93	1·3	·94	1·4	·95	1·4	·97	1·4	·98	1·4	0·99	1·4
1·48	·86	1·2	·87	1·2	·87	1·3	·88	1·3	·90	1·3	·91	1·3	·92	1·4	·93	1·4	·94	1·4	·96	1·4	·97	1·4	·98	1·5
1·50	0·85	1·2	0·86	1·3	0·87	1·3	0·88	1·3	0·89	1·3	0·90	1·4	0·91	1·4	0·92	1·4	0·93	1·4	0·95	1·4	0·96	1·5	0·97	1·5
1·52	·84	1·3	·85	1·3	·86	1·3	·87	1·3	·88	1·3	·90	1·4	·90	1·4	·91	1·4	·93	1·4	·94	1·5	·95	1·5	·96	1·5
1·54	·84	1·3	·84	1·3	·85	1·3	·86	1·3	·87	1·4	·89	1·4	·90	1·4	·91	1·4	·92	1·5	·93	1·5	·94	1·5	·95	1·5
1·56	·83	1·3	·84	1·3	·85	1·3	·85	1·4	·86	1·4	·88	1·4	·89	1·4	·90	1·5	·91	1·5	·92	1·5	·93	1·5	·94	1·6
1·58	·82	1·3	·83	1·3	·84	1·4	·85	1·4	·85	1·4	·87	1·4	·88	1·5	·89	1·5	·90	1·5	·91	1·5	·92	1·6	·93	1·6
1·60	0·82	1·3	0·82	1·3	0·83	1·4	0·84	1·4	0·85	1·4	0·86	1·5	0·87	1·5	0·88	1·5	0·89	1·5	0·90	1·5	0·91	1·6	0·92	1·6
1·62	·81	1·3	·81	1·4	·82	1·4	·83	1·4	·84	1·5	·85	1·5	·86	1·5	·87	1·5	·88	1·5	·89	1·6	·90	1·6	·91	1·6
1·64	·81	1·4	·81	1·4	·82	1·4	·83	1·4	·84	1·5	·85	1·5	·85	1·5	·86	1·5	·87	1·6	·88	1·6	·89	1·6	·90	1·6
1·66	·80	1·4	·80	1·4	·81	1·4	·82	1·5	·83	1·5	·84	1·5	·85	1·5	·85	1·6	·86	1·6	·87	1·6	·88	1·6	·90	1·7
1·68	·79	1·4	·80	1·4	·80	1·4	·81	1·5	·82	1·5	·83	1·5	·84	1·6	·84	1·6	·85	1·6	·86	1·6	·87	1·7	·89	1·7
1·70	0·79	1·4	0·79	1·4	0·80	1·5	0·81	1·5	0·81	1·5	0·83	1·6	0·83	1·6	0·84	1·6	0·85	1·6	0·86	1·7	0·87	1·7	0·88	1·7
1·72	·78	1·4	·79	1·5	·79	1·5	·80	1·5	·81	1·6	·82	1·6	·83	1·6	·83	1·6	·84	1·7	·85	1·7	·86	1·7	·87	1·7
1·74	·78	1·5	·78	1·5	·79	1·5	·79	1·5	·80	1·6	·81	1·6	·82	1·6	·82	1·7	·83	1·7	·84	1·7	·85	1·7	·86	1·8
1·76	·77	1·5	·77	1·5	·78	1·5	·78	1·6	·79	1·6	·81	1·6	·81	1·7	·82	1·7	·83	1·7	·84	1·7	·85	1·7	·86	1·8
1·78	·76	1·5	·77	1·5	·77	1·5	·78	1·6	·79	1·6	·80	1·7	·81	1·7	·81	1·7	·82	1·7	·83	1·7	·84	1·8	·85	1·8
1·80	0·76	1·5	0·77	1·5	0·77	1·6	0·78	1·6	0·78	1·6	0·79	1·7	0·80	1·7	0·81	1·7	0·81	1·7	0·82	1·8	0·83	1·8	0·84	1·8
1·82	·75	1·5	·76	1·6	·76	1·6	·77	1·6	·78	1·7	·78	1·7	·79	1·7	·80	1·7	·81	1·8	·81	1·8	·82	1·8	·83	1·8
1·84	·75	1·6	·76	1·6	·76	1·6	·77	1·6	·77	1·7	·78	1·7	·79	1·7	·79	1·8	·80	1·8	·81	1·8	·82	1·8	·83	1·9
1·86	·75	1·6	·75	1·6	·75	1·6	·76	1·7	·77	1·7	·78	1·7	·78	1·8	·79	1·8	·79	1·8	·80	1·8	·81	1·9	·82	1·9
1·88	·74	1·6	·74	1·6	·75	1·7	·75	1·7	·76	1·7	·77	1·8	·77	1·8	·78	1·8	·79	1·8	·79	1·9	·80	1·9	·81	1·9
1·90	0·74	1·6	0·74	1·6	0·74	1·7	0·75	1·7	0·76	1·7	0·76	1·8	0·77	1·8	0·78	1·8	0·78	1·8	0·79	1·9	0·80	1·9	0·81	1·9
1·92	·73	1·6	·74	1·7	·74	1·7	·75	1·7	·75	1·8	·76	1·8	·77	1·8	·77	1·8	·78	1·9	·78	1·9	·79	1·9	·80	2·0
1·94	·73	1·7	·73	1·7	·74	1·7	·74	1·7	·75	1·8	·75	1·8	·76	1·8	·76	1·9	·77	1·9	·78	1·9	·79	2·0	·79	2·0
1·96	·73	1·7	·73	1·7	·73	1·7	·74	1·7	·74	1·8	·75	1·8	·76	1·9	·76	1·9	·77	1·9	·77	1·9	·78	2·0	·79	2·0
1·98	·72	1·7	·72	1·7	·73	1·8	·73	1·8	·73	1·8	·74	1·9	·75	1·9	·75	1·9	·76	1·9	·77	2·0	·77	2·0	·78	2·0
2·00	0·72	1·7	0·72	1·7	0·72	1·8	0·73	1·8	0·73	1·8	0·74	1·9	0·74	1·9	0·75	1·9	0·75	2·0	0·76	2·0	0·77	2·0	0·78	2·1
γ →	3·7		3·8		3·9		4·0		4·1		4·2		4·25		4·3		4·35		4·4		4·45		4·5	

REFERENCES

BERKSON, J. (1953). A statistically precise and relatively simple method of estimating the bio-assay with quantal response, based on the logistic function. *J. Amer. Statist. Ass.* **48**, 565–99.

BERKSON, J. (1955). Maximum likelihood and minimum χ^2 estimates of the logistic function. *J. Amer. Statist. Ass.* **50**, 130–62.

BERKSON, J. (1957). Tables for the maximum likelihood estimate of the logistic function. *Biometrics*, **13**, 28–34.

CORNFIELD, J. & MANTEL, N. (1950). Some new aspects of the application of maximum likelihood to the calculation of the dosage response curve. *J. Amer. Statist. Ass.* **45**, 181–210.

FISHER, R. A. (1925). Theory of statistical estimation. *Proc. Camb. Phil. Soc.* **22**, 700–25.

HODGES, J. L. Jr. (1958). Fitting the logistic by maximum likelihood. *Biometrics*, **14**, 453–61.

NEYMAN, J. (1935). Su un teorema concernente le cosiddette statistiche sufficienti. *C. Ist. Ital. Attuari*, Anno VI, **13**, 3–17.

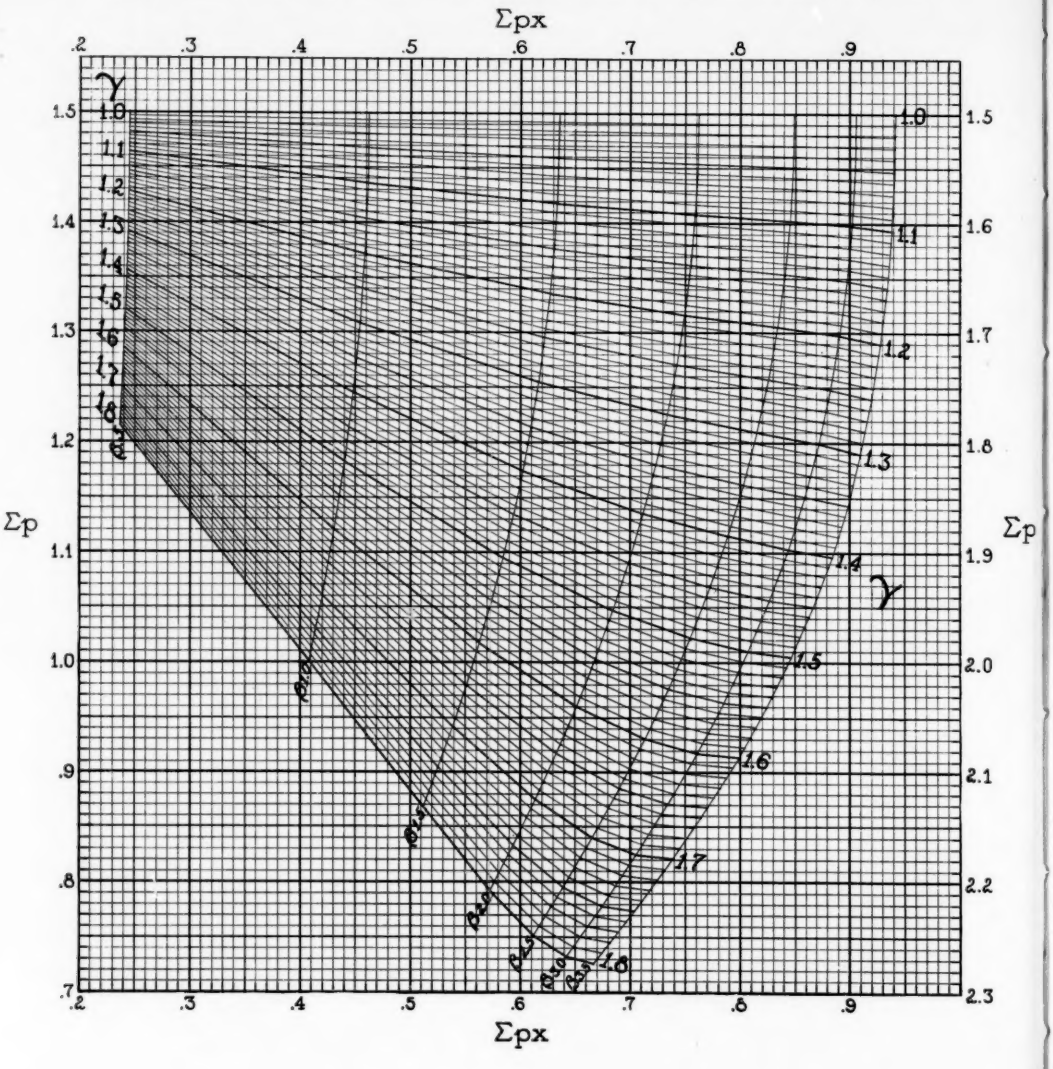
WORCESTER, J. & WILSON, E. B. (1943). A table determining L.D. 50 or the fifty per cent endpoint. *Proc. Nat. Acad. Sci., Wash.*, **29**, 207–12.

JOSEPH BERKSON

NOMOGRAMS FOR FITTING THE LOGISTIC FUNCTION

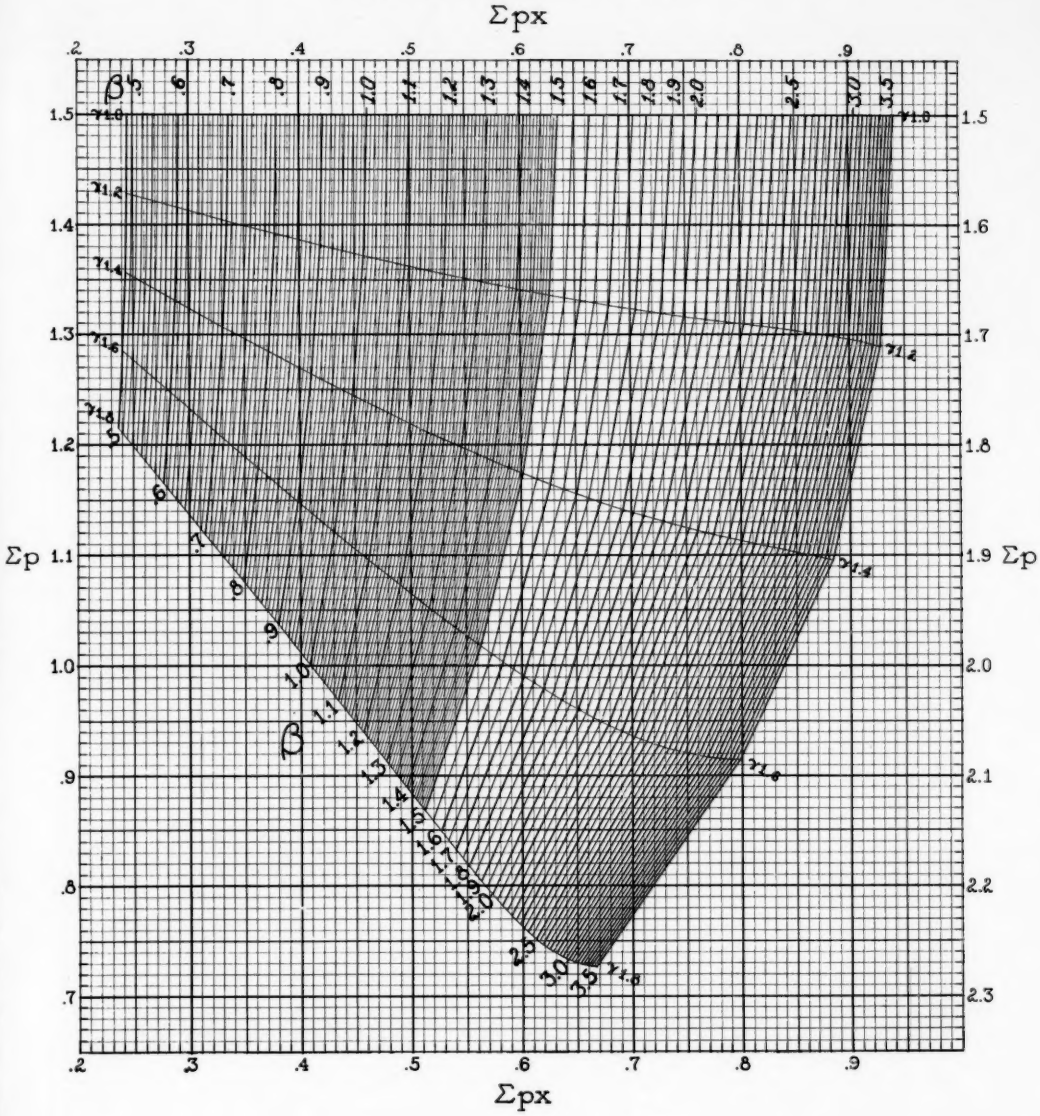
γ, (ED50), 3 doses

(For X scaled 0, 1, 2; $\Sigma px = \Sigma pX - \Sigma p$. If Σp is read on the right-hand scale, $\hat{\gamma} = 2 - \gamma$.)



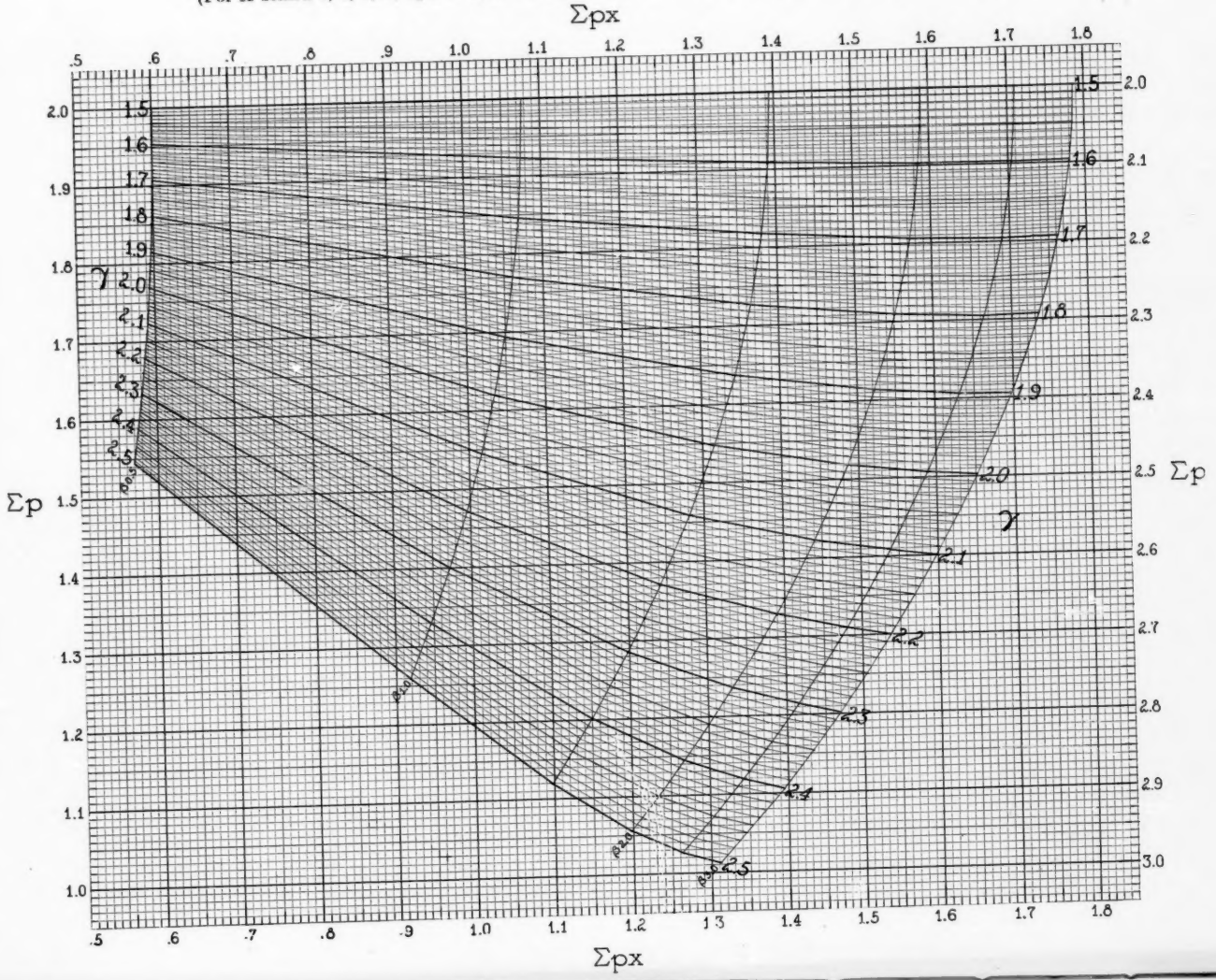
β, 3 doses

(For X scaled 0, 1, 2; $\Sigma px = \Sigma pX - \Sigma p$.)



γ, (ED 50), 4 doses

(For X scaled 0, 1, 2, 3; $\Sigma px = \Sigma pX - 1{\cdot}5\Sigma p$. If Σp is read on the right-hand scale, $\hat{\gamma} = 3 - \gamma$.)

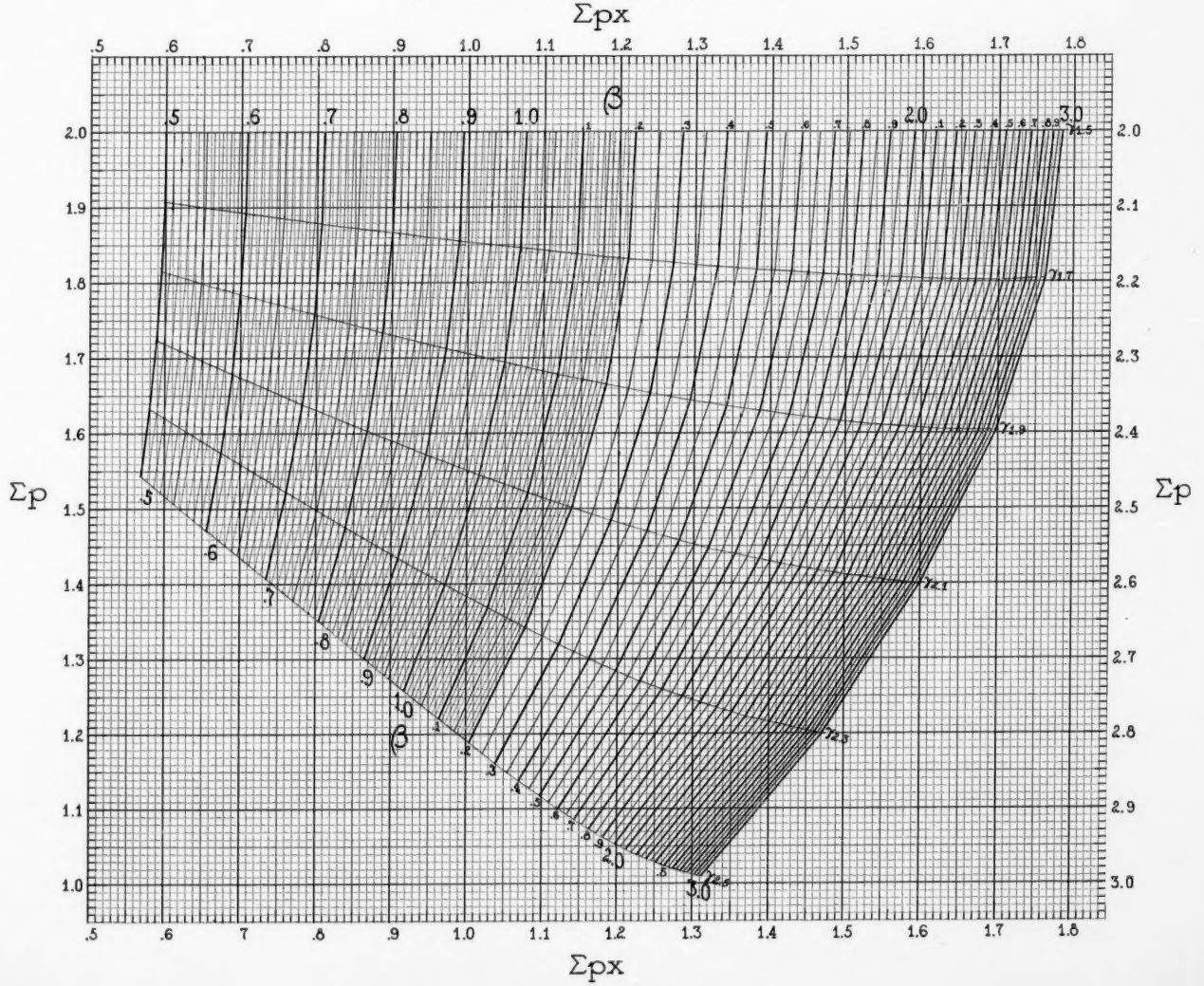


β, 4 doses

(For X scaled 0, 1, 2, 3; $\Sigma px = \Sigma pX - 1{\cdot}5\Sigma p$.)

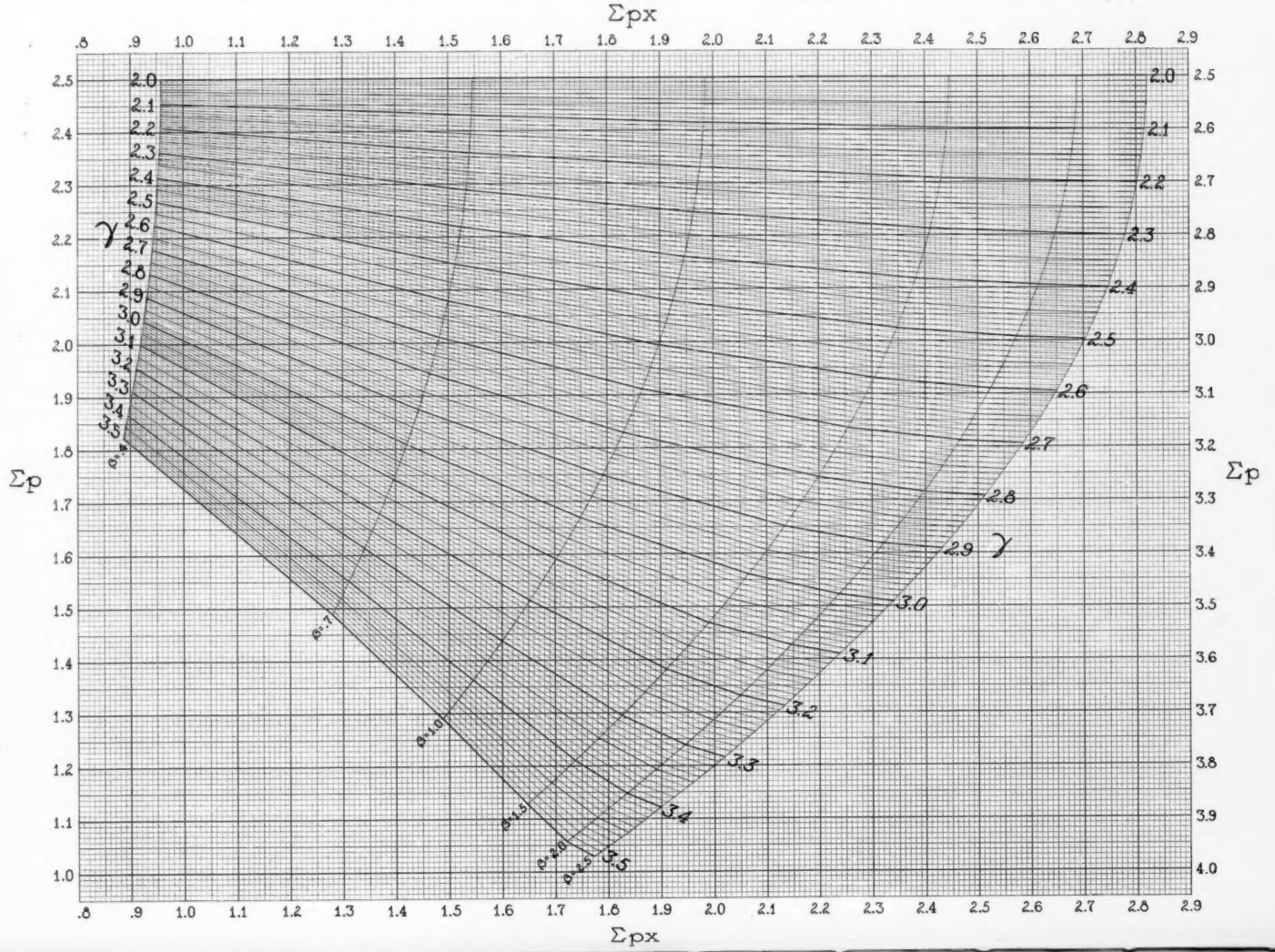
β, 4 doses

(For X scaled 0, 1, 2, 3; $\Sigma px = \Sigma pX - 1{\cdot}5\Sigma p$.)



γ, (ED 50), 5 doses

(For X scaled 0, 1, 2, 3, 4; $\Sigma px = \Sigma pX - 2\Sigma p$. If Σp is read on the right-hand scale, $\hat{\gamma} = 4 - \gamma$.)

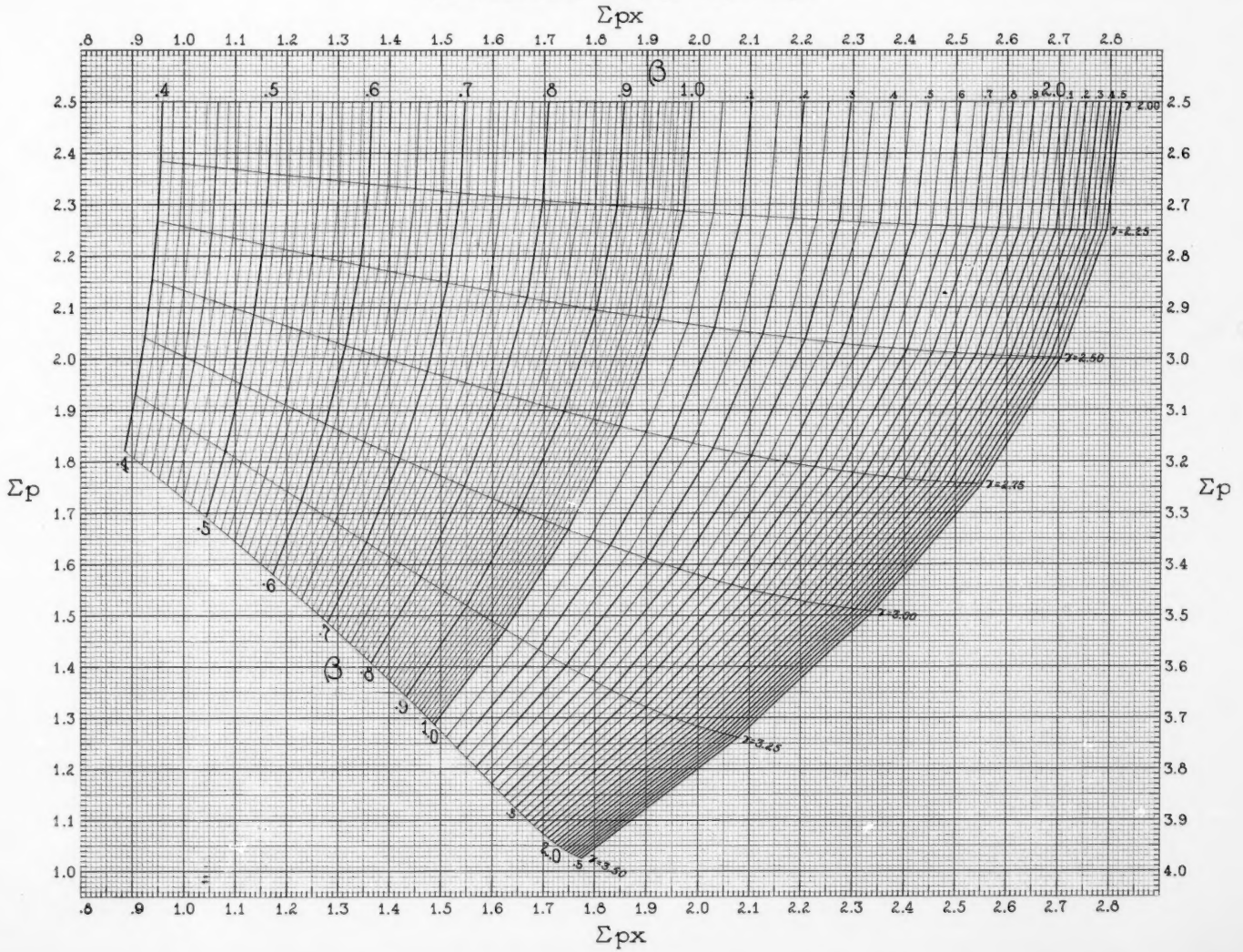


β, 5 doses

(For X scaled 0, 1, 2, 3, 4; $\Sigma px = \Sigma pX - 2\Sigma p$.)

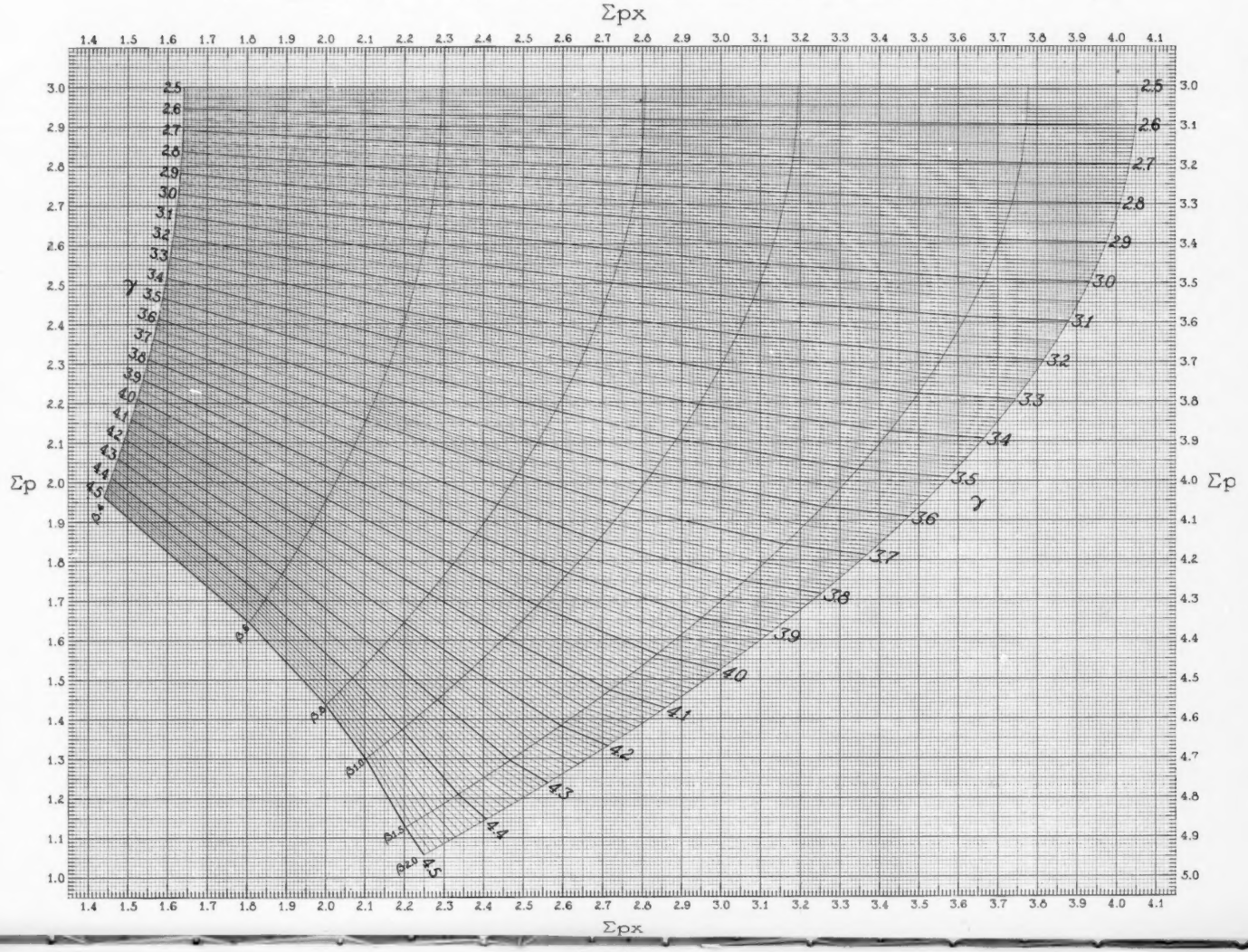
β, 5 doses

(For X scaled 0, 1, 2, 3, 4; $\Sigma px = \Sigma pX - 2\Sigma p$.)



γ, (ED 50), 6 doses

(For X scaled 0, 1, 2, 3, 4, 5; $\Sigma px = \Sigma pX - 2{\cdot}5\Sigma p$. If Σp is read on the right-hand scale, $\hat{\gamma} = 5 - \gamma$.)

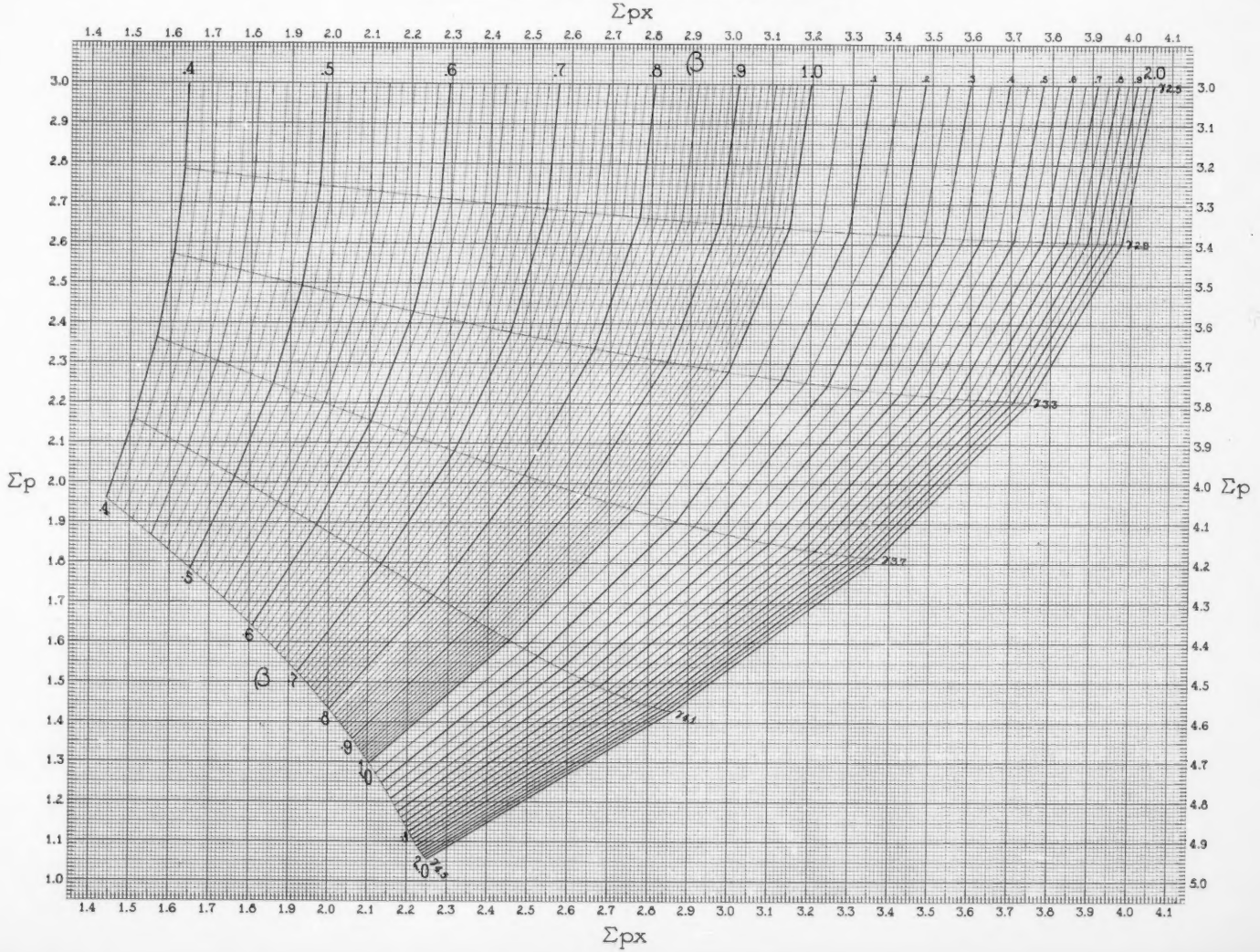


β, 6 doses

(For X scaled 0, 1, 2, 3, 4, 5; $\Sigma px = \Sigma pX - 2{\cdot}5\Sigma p$.)

β, 6 doses

(For X scaled 0, 1, 2, 3, 4, 5; $\Sigma px = \Sigma pX - 2{\cdot}5\Sigma p$.)



Biometrika (1960), **47**, 1 *and* 2, *p.* 143
Printed in Great Britain

The Borel–Tanner distribution

BY FRANK A. HAIGHT AND MELVIN ALLEN BREUER
Institute of Transportation and Traffic Engineering, University of California, Los Angeles

1. INTRODUCTION

Let $\alpha > 0$ and r be a positive integer. Then

$$p(x; r, \alpha) = A(x, r)\, e^{-\alpha x}\, \alpha^{x-r} \quad (x = r, r+1, \ldots) \tag{1}$$

is a probability distribution when

$$A(x, r) = \frac{r}{(x-r)!}\, x^{x-r-1}. \tag{2}$$

Borel (1942) first showed this for the case $r = 1$ by a geometric argument, and Tanner (1953) pointed out that Borel's argument is in fact valid for general r. In the theory of queues $p(x; r, \alpha)$ represents the probability that exactly x members of a queue will be served before the queue first vanishes, beginning with r members, and with α equal to the traffic intensity, assuming Poisson arrivals and constant service time. A continuous analogue of this distribution occurs in a paper by Kendall (1957), p. 211, equation (20).

Another interesting property of the distribution is its domain of definition. Most discrete distributions are defined over some finite subset of the integers (e.g. binomial, hypergeometric) or over all of the non-negative integers (e.g. Poisson, negative binomial). In cases where such a domain is inapplicable, it is necessary to truncate the domain; for example, the Poisson is often used without $x = 0$. The Borel–Tanner distribution, however, is available for any positive integral starting point.

First, we suppose that $\alpha < 1$, and let

$$\beta = \alpha\, e^{-\alpha}. \tag{3}$$

Then

$$\beta^{x-r}\, e^{-\alpha r} = \alpha^{x-r}\, e^{-\alpha x} \tag{4}$$

and

$$\frac{d\alpha}{d\beta} = \frac{\alpha}{\beta(1-\alpha)} \tag{5}$$

and

$$\frac{d^2\alpha}{d\beta^2} = \frac{(2-\alpha)\,\alpha^2}{(1-\alpha)^3\,\beta^2}. \tag{6}$$

Using these formulae, the mean may be calculated as follows

$$\left.\begin{aligned} \mu &= \Sigma x p(x; r, \alpha), \\ \alpha^r \mu &= \Sigma x A(x, r)\, \beta^x, \end{aligned}\right\} \tag{7}$$

where these (and all subsequent) summations are over $x = r, r+1, \ldots$. Substituting (4) into

$$1 = \Sigma A(x, r)\, \beta^{x-r}\, e^{-\alpha r}$$

gives

$$\alpha^r = \Sigma A(x, r)\, \beta^x \tag{8}$$

and differentiating with respect to α, we have

$$\beta r\alpha^{r-1}\frac{d\alpha}{d\beta} = \Sigma A(x,r)\,x\beta^x. \tag{9}$$

From (5), (7) and (9), we obtain $$\mu = \frac{r}{1-\alpha}. \tag{10}$$

This result was obtained by Borel for $r = 1$. The variance can be found similarly

$$\mu_2 = \sigma^2 + \mu^2 = \Sigma x^2 p(x; r, \alpha).$$

Differentiating (9) and using (6) gives

$$\sigma^2 = \frac{\alpha r}{(1-\alpha)^3}. \tag{11}$$

Let $\phi(s)$ be the probability generating function of the Borel–Tanner distribution. Then

$$\alpha^r \phi(s) = \Sigma(s\beta)^x A(x,r) \quad (0 < s < 1),$$

and substituting $s\beta$ for β in (8), we find

$$\phi(s) = [\alpha(s\beta)/\alpha(\beta)]^r, \tag{12}$$

where $\alpha(\beta)$ denotes the functional dependence defined by (3). Using (3), we can write

$$\phi(s) = s^r \exp[r\alpha(s\beta) - r\alpha(\beta)]. \tag{13}$$

Since $\phi(1) = 1$, $\phi'(1) = \mu$, $\phi''(1) = \sigma^2 + \mu^2 - \mu$, (10) and (11) can be confirmed by differentiation of (13).

The function $\alpha(\beta)$ can be written explicitly by use of Lagrange's Theorem (cf. Whittaker & Watson, 1935)

$$\alpha = \sum_1^\infty \frac{n^{n-1}}{n!}\beta^n. \tag{14}$$

Then $\phi(s)$ can be written in terms of s and the parameters r and α

$$\phi(s) = s^r \exp r\left[\sum_1^\infty \frac{n^{n-1}\alpha^n}{n!\,e^{\alpha n}}(s^n - 1)\right]. \tag{15}$$

Differentiating this with respect to s and using (10) when $s = 1$, we obtain

$$\sum_1^\infty \frac{n^n\alpha^n}{n!\,e^{\alpha n}} = \frac{\alpha}{1-\alpha}. \tag{16}$$

The relationship between s and ϕ is fairly complicated in both (12) and (15). Two simpler formulae are

$$\log\phi - \alpha r\phi^{1/r} - r\log s + \alpha r = 0 \tag{17}$$

and, using subscripts to denote partial differentiation,

$$\alpha\phi_\alpha - s\phi_s + \alpha s\phi_s + r\phi = 0. \tag{18}$$

If $K(s)$ denotes the cumulant generating function and $M(s)$ the moment generating function, then (17) is equivalent to

$$K(s) = \alpha r[M(s)]^{1/r} + rs - \alpha r \tag{19}$$

and higher moments can be obtained fairly easily by differentiation of this formula.

Turning now to the case $\alpha > 1$, we will show that $\Sigma p(x; r, \alpha) < 1$, and thereby give a method for computing $p(\infty; r, \alpha) = 1 - \Sigma p(x; r, \alpha)$. Let α' denote the smaller root of (3), which is given by equation (14). Then

$$\Sigma p(x; r, \alpha) = \Sigma A(x, r)\, \beta^{x-r} e^{-\alpha x}$$

$$= e^{(\alpha' - \alpha) r} \tag{20}$$

and therefore

$$p(\infty; r, \alpha) = 1 - e^{(\alpha' - \alpha) r} = 1 - (\alpha'/\alpha)^r, \tag{21}$$

where β is computed from (3) and then α' from (14). In the intermediate case $\alpha = 1$, we have $\alpha' = \alpha$, and therefore the infinite component is zero.

Finally, we mentioned two additional properties of the Borel–Tanner distribution. (*a*) If r is fixed, maximum likelihood estimation of α is equivalent to the method of moments, i.e.

$$\hat{\alpha} = \frac{\bar{x} - r}{\bar{x}}. \tag{22}$$

(*b*) Since the Borel–Tanner distribution has a very long tail, the modal value M is of some interest. If M is large enough to justify the use of Stirling's approximation, and $\alpha \leqslant 1$, then

$$f^{-1}(\beta e) - 1 < M < f^{-1}(\beta e), \tag{23}$$

where

$$f(x) = \frac{(x - r)\, x^{r + \frac{1}{2}}}{(x - 1)^{r + \frac{3}{2}}}. \tag{24}$$

$f(x)$ is increasing from $x = r$ to $x = \frac{2}{3}r^2 + 3r$, at which value $f > 1$. However, $f(M) < 1$, and therefore the values of $f(x)$ are only needed from $x = r + 1$ to the first argument yielding $f(x) \geqslant 1$ for determination of M.

When $r = 1$, $M = 1$, and when $\alpha > 1$, the greatest finite ordinate is r.

2. Tables

In the table which follows, we give values of the Borel–Tanner distribution, correct to five places of decimals, in cumulative form, i.e.

$$P(x; r, \alpha) = \sum_{y=r}^{x} p(y; r, \alpha),$$

for the parameter values $r = 1$, $\alpha = 0{\cdot}01\,(0{\cdot}01)\,0{\cdot}62$. In every case, the values of x taken are large enough to ensure that $P(x; r, \alpha) > 0{\cdot}999$.

Certain additional tables, some of which were required by the authors in a special investigation have been computed, but owing to their incomplete character they are not reproduced here. They are summarized under headings I–V.

I. Values of $P(x; r, \alpha)$ for $r = 1$, $\alpha = 0{\cdot}63\,(0{\cdot}01)\,0{\cdot}99$, with the x values arbitrarily stopped at $x = 34$. In these cases, the cumulative sums fall short of 0·999; typical final values are

$$\alpha = 0{\cdot}7, \quad P = 0{\cdot}99562; \qquad \alpha = 0{\cdot}8, \quad P = 0{\cdot}97911,$$

$$\alpha = 0{\cdot}9, \quad P = 0{\cdot}93711; \qquad \alpha = 0{\cdot}99, \quad P = 0{\cdot}87280.$$

II. Values of $p(x; r, \alpha)$, correct to five decimal places, with

α	r	α	r	α	r
0·1, 0·2, 0·3, 0·4	1 (1) 30	0·5	1 (1) 20	0·9	1 (1) 8
		0·6	1 (1) 24	0·99, 0·999	1 (1) 5
		0·7	1 (1) 19		
		0·8	1 (1) 10		

In each of these cases, the value of x is taken up to the point where $p(x; r, \alpha) < 0{\cdot}01$, but never beyond $x = 51$. Owing to the very long tail of the Borel–Tanner distribution, this frequently means that $P(x; r, \alpha)$ is substantially less than unity.

III. Values of $P(x; r, \alpha)$, correct to four decimal places, for $r = 1$, $\alpha = 1{\cdot}0\,(0{\cdot}1)\,2\,(1)\,10$, including the infinite value given by equation (21). In the majority of cases, values of x are great enough to ensure that $P(x; r, \alpha) > 0{\cdot}99$, but for $\alpha = 1,\ 1{\cdot}1$, the sequence has been truncated at $x = 73$.

IV. For calculation of the mode, tables of $f(x)$ from equation (24) for $r = 2\,(1)\,20$ with x going from r to a value for which $f(x) > 1$.

V. Fragmentary tables of $p(x; r, \alpha)$ and $P(x; r, \alpha)$, without the infinite value, but with x going in many cases to very large values (> 100), and with the following parameter values:

α	r
1·0, 1·3 (0·1) 1·9	2, 3, 4, 5
1·1, 1·2	2 (1) 10
2	2, 3, 4, 5
3	2, 3, 4
4, 5	2

We invite readers' suggestions for further calculation or publication of Borel–Tanner tables, and in the meantime shall be glad to make our unpublished material available to anyone wishing to use any special results from among those listed above.

REFERENCES

BOREL, E. (1942). Sur l'emploi du théorème de Bernoulli pour faciliter le calcul d'un infinité de coefficients. Application au problème de l'attente à un guichet. *C.R. Acad. Sci., Paris*, **214**, 452–6.

KENDALL, D. G. (1957). Some problems in the theory of dams. *J. Roy. Statist. Soc.* B, **19**, 207–33.

TANNER, J. C. (1953). A problem of interference between two queues. *Biometrika*, **40**, 58–69.

WHITTAKER, E. T. & WATSON, G. N. (1935). *A Course of Modern Analysis* (4th edition). Cambridge University Press.

Table 1. *The Borel–Tanner distribution* ($r = 1$)

x	$P(x; 1, \alpha)$
	$\alpha = 0{\cdot}01$
1	0·99905
2	·99985
	$\alpha = 0{\cdot}02$
1	0·98020
2	·99941
	$\alpha = 0{\cdot}03$
1	0·97045
2	·99870
3	·99993
	$\alpha = 0{\cdot}04$
1	0·96079
2	·99771
3	·99984
	$\alpha = 0{\cdot}05$
1	0·95123
2	·99647
3	·99970
	$\alpha = 0{\cdot}06$
1	0·94176
2	·99498
3	·99949
	$\alpha = 0{\cdot}07$
1	0·93239
2	·99325
3	·99921
	$\alpha = 0{\cdot}08$
1	0·92312
2	·99129
3	·99884
4	·99983
	$\alpha = 0{\cdot}09$
1	0·91393
2	·98911
3	·99838
4	·99974
	$\alpha = 0{\cdot}10$
1	0·90484
2	·98671
3	·99782
4	·99961

x	$P(x; 1, \alpha)$
	$\alpha = 0{\cdot}11$
1	0·89583
2	·98411
3	·99716
4	·99945
	$\alpha = 0{\cdot}12$
1	0·88692
2	·98132
3	·99639
4	·99924
	$\alpha = 0{\cdot}13$
1	0·87810
2	·97833
3	·99550
4	·99898
5	·99976
	$\alpha = 0{\cdot}14$
1	0·86936
2	·97517
3	·99449
4	·99866
5	·99966
	$\alpha = 0{\cdot}15$
1	0·86071
2	·97183
3	·99335
4	·99829
5	·99954
	$\alpha = 0{\cdot}16$
1	0·85214
2	·96833
3	·99209
4	·99785
5	·99938
	$\alpha = 0{\cdot}17$
1	0·84366
2	·96467
3	·99070
4	·99733
5	·99919
	$\alpha = 0{\cdot}18$
1	0·83527
2	·96085
3	·98917

x	$P(x; 1, \alpha)$
	$\alpha = 0{\cdot}18$ (*cont.*)
4	0·99674
5	·99897
6	0·99966
	$\alpha = 0{\cdot}19$
1	0·82696
2	·95689
3	·98752
4	·99607
5	·99869
6	0·99955
	$\alpha = 0{\cdot}20$
1	0·81873
2	·95279
3	·98572
4	·99531
5	·99837
6	0·99942
	$\alpha = 0{\cdot}21$
1	0·81058
2	·94856
3	·98379
4	·99446
5	·99800
6	0·99925
	$\alpha = 0{\cdot}22$
1	0·80252
2	·94421
3	·98173
4	·99351
5	·99757
6	0·99906
	$\alpha = 0{\cdot}23$
1	0·79453
2	·93973
3	·97953
4	·99246
5	·99707
6	0·99882
7	·99951
	$\alpha = 0{\cdot}24$
1	0·78663
2	·93514

x	$P(x; 1, \alpha)$
	$\alpha = 0{\cdot}24$ (*cont.*)
3	0·97719
4	·99131
5	·99651
6	0·99855
7	·99938
	$\alpha = 0{\cdot}25$
1	0·77880
2	·93043
3	·97472
4	·99005
5	·99588
6	0·99823
7	·99922
	$\alpha = 0{\cdot}26$
1	0·77105
2	·92563
3	·97211
4	·98868
5	·99516
6	0·99786
7	·99903
	$\alpha = 0{\cdot}27$
1	0·76338
2	·92072
3	·96937
4	·98719
5	·99437
6	0·99743
7	·99880
8	·99943
	$\alpha = 0{\cdot}28$
1	0·75578
2	·91572
3	·96649
4	·98559
5	·99349
6	0·99695
7	·99853
8	·99928
	$\alpha = 0{\cdot}29$
1	0·74826
2	·91063
3	·96348
4	·98387

x	$P(x; 1, \alpha)$
	$\alpha = 0{\cdot}29$ (*cont.*)
5	0·99251
6	0·99640
7	·99823
8	·99911
	$\alpha = 0{\cdot}30$
1	0·74082
2	·90546
3	·96035
4	·98203
5	·99145
6	0·99579
7	·99787
8	·99890
9	·99942
	$\alpha = 0{\cdot}31$
1	0·73345
2	·90021
3	·95708
4	·98007
5	·99028
6	0·99510
7	·99746
8	·99866
9	·99928
	$\alpha = 0{\cdot}32$
1	0·72615
2	·89488
3	·95369
4	·97799
5	·98902
6	0·99433
7	·99700
8	·99838
9	·99911
	$\alpha = 0{\cdot}33$
1	0·71892
2	·88948
3	·95018
4	·97578
5	·98764
6	0·99348
7	·99647
8	·99805
9	·99891
10	·99938

Table 1 (*cont.*)

x	$P(x; 1, \alpha)$
$\alpha = 0{\cdot}34$	
1	0·71177
2	·88402
3	·94655
4	·97345
5	·98616
6	0·99254
7	·99588
8	·99768
9	·99867
10	·99923
$\alpha = 0{\cdot}35$	
1	0·70469
2	·87849
3	·94279
4	·97099
5	·98457
6	0·99152
7	·99522
8	·99725
9	·99840
10	·99905
$\alpha = 0{\cdot}36$	
1	0·69768
2	·87291
3	·93892
4	·96840
5	·98286
6	0·99039
7	·99448
8	·99677
9	·99808
10	·99885
11	0·99930
$\alpha = 0{\cdot}37$	
1	0·69073
2	·86727
3	·93494
4	·96569
5	·98104
6	0·98917
7	·99366
8	·99622
9	·99771
10	·99860
11	0·99913

x	$P(x; 1, \alpha)$
$\alpha = 0{\cdot}38$	
1	0·68386
2	·86157
3	·93085
4	·96285
5	·97909
6	0·98785
7	·99276
8	·99561
9	·99730
10	·99832
11	0·99894
12	·99933
$\alpha = 0{\cdot}39$	
1	0·67706
2	·85584
3	·92665
4	·95989
5	·97703
6	0·98641
7	·99177
8	·99492
9	·99682
10	·99799
11	0·99871
12	·99917
$\alpha = 0{\cdot}40$	
1	0·67032
2	·85005
3	·92234
4	·95680
5	·97484
6	0·98487
7	·99069
8	·99416
9	·99628
10	·99761
11	0·99844
12	·99898
13	·99933
$\alpha = 0{\cdot}41$	
1	0·66365
2	·84423
3	·91793
4	·95358
5	·97253

x	$P(x; 1, \alpha)$
$\alpha = 0{\cdot}41$ (*cont.*)	
6	0·98322
7	·98950
8	·99332
9	·99568
10	·99718
11	0·99814
12	·99876
13	·99917
$\alpha = 0{\cdot}42$	
1	0·65705
2	·83837
3	·91342
4	·95024
5	·97009
6	0·98144
7	·98822
8	·99238
9	·99500
10	·99669
11	0·99778
12	·99850
13	·99898
14	·99930
$\alpha = 0{\cdot}43$	
1	0·65051
2	·83247
3	·90882
4	·94678
5	·96752
6	·097955
7	·98683
8	·99136
9	·99425
10	·99613
11	0·99737
12	·99820
13	·99876
14	·99914
$\alpha = 0{\cdot}44$	
1	0·64404
2	·82654
3	·90412
4	·94320
5	·96483
6	0·97754
7	·98532

x	$P(x; 1, \alpha)$
$\alpha = 0{\cdot}44$ (*cont.*)	
8	0·99024
9	·99342
10	·99551
11	0·99690
12	·99785
13	·99850
14	·99894
15	·99925
$\alpha = 0{\cdot}45$	
1	0·63763
2	·82058
3	·89933
4	·93950
5	·96201
6	0·97540
7	·98371
8	·98902
9	·99249
10	·99481
11	0·99638
12	·99745
13	·99819
14	·99871
15	·99908
$\alpha = 0{\cdot}46$	
1	0·63128
2	·81460
3	·89445
4	·93568
5	·95906
6	0·97314
7	·98197
8	·98769
9	·99148
10	·99403
11	0·99578
12	·99699
13	·99784
14	·99844
15	·99887
16	0·99918
$\alpha = 0{\cdot}47$	
1	0·62500
2	·80860
3	·88950
4	·93174
5	·95598

x	$P(x; 1, \alpha)$
$\alpha = 0{\cdot}47$ (*cont.*)	
6	0·97074
7	·98012
8	·98625
9	·99036
10	·99317
11	0·99511
12	·99647
13	·99744
14	·99813
15	·99862
16	0·99899
17	·99925
$\alpha = 0{\cdot}48$	
1	0·61878
2	·80257
3	·88445
4	·92769
5	·95277
6	0·96822
7	·97814
8	·98470
9	·98915
10	·99221
11	0·99436
12	·99588
13	·99698
14	·99776
15	·99834
16	0·99876
17	·99907
$\alpha = 0{\cdot}49$	
1	0·61263
2	·79653
3	·87934
4	·92353
5	·94944
6	0·96557
7	·97603
8	·98303
9	·98782
10	·99116
11	0·99353
12	·99522
13	·99645
14	·99735
15	·99801

Table 1 (*cont.*)

x	$P(x; 1, \alpha)$
$\alpha = 0{\cdot}49$ (*cont.*)	
16	0·99850
17	·99886
18	·99913
$\alpha = 0{\cdot}50$	
1	0·60653
2	·79047
3	·87414
4	·91926
5	·94598
6	0·96278
7	·97379
8	·98124
9	·98638
10	·99001
11	0·99260
12	·99448
13	·99585
14	·99687
15	·99762
16	0·99819
17	·99861
18	·99893
19	·99918
$\alpha = 0{\cdot}51$	
1	0·60050
2	·78440
3	·86888
4	·91488
5	·94239
6	0·95986
7	·97143
8	·97932
9	·98483
10	·98875
11	0·99158
12	·99365
13	·99518
14	·99632
15	·99718
16	0·99783
17	·99832
18	·99869
19	·99898
20	·99920
$\alpha = 0{\cdot}52$	
1	0·59452
2	·77832
3	0·86355
4	·91039
5	·93868
6	0·95681
7	·96892
8	·97727
9	·98315
10	·98738
11	0·99046
12	·99273
13	·99443
14	·99570
15	·99667
16	0·99741
17	·99797
18	·99841
19	·99875
20	·99901
$\alpha = 0{\cdot}53$	
1	0·58860
2	·77223
3	·85815
4	·90580
5	·93484
6	0·95362
7	·96629
8	·97509
9	·98135
10	·98589
11	0·98923
12	·99172
13	·99359
14	·99501
15	·99609
16	0·99693
17	·99758
18	·99808
19	·99848
20	·99879
21	0·99903
$\alpha = 0{\cdot}54$	
1	0·58275
2	·76613
3	·85269
4	·90112
5	·93088
6	0·95030
7	·96351
8	·97277
9	·97942
10	·98428
11	0·98788
12	·99059
13	·99265
14	·99422
15	·99544
16	0·99638
17	·99712
18	·99770
19	·99816
20	·99852
21	0·99880
22	·99903
$\alpha = 0{\cdot}55$	
1	0·57695
2	·76003
3	·84717
4	·89633
5	·92680
6	0·94685
7	·96060
8	·97032
9	·97735
10	·98254
11	0·98642
12	·98936
13	·99161
14	·99335
15	·99470
16	0·99576
17	·99659
18	·99725
19	·99778
20	·99820
21	0·99854
22	·99881
23	·99903
$\alpha = 0{\cdot}56$	
1	0·57121
2	·75393
3	·84160
4	·89145
5	·92260
6	0·94326
7	·95754
8	·96772
9	·97515
10	·98067
11	0·98483
12	·98801
13	·99047
14	·99237
15	·99387
16	0·99505
17	·99599
18	·99674
19	·99734
20	·99783
21	0·99822
22	·99854
23	·99879
24	·99901
$\alpha = 0{\cdot}57$	
1	0·56553
2	·74782
3	·83597
4	·88648
5	·91828
6	0·93954
7	·95435
8	·96499
9	·97281
10	·97867
11	0·98312
12	·98654
13	·98921
14	·99129
15	·99295
16	0·99426
17	·99531
18	·99616
19	·99684
20	·99740
21	0·99785
22	·99822
23	·99852
24	·99877
25	·99897
26	0·99914
$\alpha = 0{\cdot}58$	
1	0·55990
2	·74172
3	·83029
4	·88142
5	·91385
6	0·93569
7	·95102
8	·96211
9	·97032
10	·97652
11	0·98126
12	·98494
13	·98782
14	·99010
15	·99191
16	0·99337
17	·99454
18	·99549
19	·99626
20	·99689
21	0·99741
22	·99784
23	·99819
24	·99848
25	·99873
26	0·99893
27	·99910
$\alpha = 0{\cdot}59$	
1	0·55433
2	·73562
3	·82456
4	·87627
5	·90931
6	0·93171
7	·94754
8	·95908
9	·96769
10	·97423
11	0·97927
12	·98321
13	·98632
14	·98879
15	·99077
16	0·99237
17	·99367
18	·99473
19	·99560
20	·99632

Table 1 (*cont.*)

x	$P(x; 1, \alpha)$
$\alpha = 0{\cdot}59$ (*cont.*)	
21	0·99691
22	·99740
23	·99781
24	·99815
25	·99843
26	0·99867
27	·99887
28	·99904
$\alpha = 0{\cdot}60$	
1	0·54881
2	·72953
3	·81879
4	·87104
5	·90465
6	0·92760
7	·94393
8	·95591
9	·96491
10	·97179
11	0·97714
12	·98134
$\alpha = 0{\cdot}60$ (*cont.*)	
13	0·98468
14	·98735
15	·98951
16	0·99126
17	·99270
18	·99388
19	·99485
20	·99566
21	0·99633
22	·99689
23	·99736
24	·99775
25	·99809
26	0·99837
27	·99860
28	·99880
29	·99897
30	·99912
$\alpha = 0{\cdot}61$	
1	0·54335
2	·72344
$\alpha = 0{\cdot}61$ (*cont.*)	
3	0·81298
4	·86573
5	·89989
6	0·92336
7	·94017
8	·95259
9	·96198
10	·96921
11	0·97485
12	·97932
13	·98290
14	·98578
15	·98812
16	0·99004
17	·99161
18	·99292
19	·99400
20	·99491
21	0·99566
22	·99630
23	·99684
24	·99729
25	0·99767
26	0·99800
27	·99828
28	·99852
29	·99872
30	·99889
31	0·99904
$\alpha = 0{\cdot}62$	
1	0·53794
2	·71736
3	·80712
4	·86035
5	·89502
6	0·91899
7	·93629
8	·94912
9	·95889
10	·96647
11	0·97242
12	·97716
$\alpha = 0{\cdot}62$ (*cont.*)	
13	0·98097
14	·98409
15	·98660
16	0·98868
17	·99041
18	·99185
19	·99305
20	·99405
21	0·99490
22	·99562
23	·99623
24	·99675
25	·99719
26	0·99757
27	·99790
28	·99817
29	·99841
30	·99862
31	0·99880
32	·99895
33	·99908

Biometrika (1960), **47**, 1 *and* 2, *p.* 151
Printed in Great Britain

The distribution of Kendall's score S for a pair of tied rankings

BY E. J. BURR
University of New England, Armidale, N.S.W., Australia

1. INTRODUCTION AND SUMMARY

In a scatter diagram of a bivariate population (or sample), if both variates are measured on an interval scale, we say that there is linear regression or curvilinear regression if the means of arrays tend to lie close to a straight line or curve, respectively. But if either or both variates are measured only on an ordinal or ranking scale, in which the ratios of the intervals between consecutive ranks are undefined or unknown, then we can no longer distinguish between a (straight) line of regression and a monotonic curve of regression. (A rising monotonic curve is one for which $x_1 > x_2$ implies $y_1 \geqslant y_2$ where (x_1, y_1) and (x_2, y_2) are the co-ordinates of any two points on the curve. For a falling monotonic curve, $x_1 > x_2$ implies $y_1 \leqslant y_2$.) In this case the terms linear correlation and linear regression are inappropriate. We propose to describe this type of association between rankings by the term *monotonic correlation*, to distinguish it from cases where the regression curve contains both rising and falling arcs.

Given a random sample of n members ranked according to two different properties U, V, it is often required to test the null hypothesis that there is no association between U, V in the parent population, against the alternative hypothesis that there is some degree of monotonic correlation. The most convenient test statistic for this purpose is Kendall's score (or rank sum) S.

Methods for finding the sampling distributions of S for $n = 2, 3, 4, \ldots$ have been described by Kendall (1955) for the case when both rankings are free from ties, and by Sillitto (1947) for the case when one ranking is free from ties. But in the general case when both rankings contain ties, the only clues to the distribution of S up to the present time have been a knowledge of its mean and variance and a conjecture that 'there is probably little important error involved in using the normal approximation for $n \geqslant 10$, unless the ties are very extensive or very numerous' (Kendall, 1955, §4·10).

In this paper it is shown how the exact distribution of S for a sample of n members, with any specified permutations of tied groups in each ranking, may be expressed as a linear combination of certain distributions for samples of fewer than n members (§6). A table is given showing the 241 essentially distinct distributions of S for samples with $2 \leqslant n \leqslant 6$. Instructions and worked examples are supplied (§7) showing how to use this table (*a*) to test the significance of an observed score when $n \leqslant 6$, and (*b*) to compute sampling distributions when $n > 6$. Calculations of the second kind are sometimes tedious, but it is shown how they may often be avoided, either by using an independent short method for finding the exact significance levels of the two or three extreme values of S and of $-S$ (§8), or by using the normal approximation (§9). Determination of the continuity correction for use with the normal approximation is discussed, and examples with $n = 8$, 10 and 12 are given showing the magnitudes of the errors involved in the normal approximation. In the cases examined, Kendall's conjecture is found to be correct except when the ties

are so extensive that only a few of the extreme values of S are significant, in which case the normal approximation may be avoided by using the method of § 8.

2. THE RANKINGS

Let n members be ranked or classified according to the degree to which they possess a property U. Suppose for the sake of definiteness that, in the case in question, only three different degrees of U are distinguishable, so that the ranking separates the n members into three classes, having u_1 members in the first class, u_2 in the second and u_3 in the third, with $u_1+u_2+u_3 = n$. For brevity we write

$$\{u\} \equiv (u_1, u_2, u_3),$$

and we call the terms u_i of $\{u\}$ the class frequencies of the first ranking.

If every class frequency is unity, so that there are n classes, we say that the ranking is untied; otherwise, two or more members in the same class (that is, ranked equal) are said to be tied. A ranking containing only two classes is called a dichotomy. In a ranking with three or more classes, the classes must have a natural order; in other words, the level of measurement must be ordinal and not merely nominal (Siegel, 1956).

Let the same n members be ranked again according to a second property V, and suppose that in this case four different degrees of V are distinguishable. We write

$$\{v\} \equiv (v_1, v_2, v_3, v_4),$$

where the terms v_j of $\{v\}$ are the class frequencies of the second ranking, and $\Sigma v_j = n$.

Table 1. *Ordered contingency or frequency table showing two rankings of n members*

a_{11}	a_{12}	a_{13}	a_{14}	u_1
a_{21}	a_{22}	a_{23}	a_{24}	u_2
a_{31}	a_{32}	a_{33}	a_{34}	u_3
v_1	v_2	v_3	v_4	

First ranking

Second ranking

The result of the two rankings may be conveniently displayed in the form of an ordered contingency or frequency table (Table 1) with row totals $\{u\}$ and column totals $\{v\}$. We call Table 1 a 3×4 table of n members. It may be thought of as a scatter diagram for a bivariate sample. For brevity we write $\{a\}$ for the array of twelve cell frequencies or elements a_{ij}. Some numerical examples of such arrays are given in § 8.

In the sequel we shall often make statements about a 3×4 table, but it is to be understood that the numbers 3, 4 have no special significance, and are chosen only for convenience of exposition. Such statements will be generally true (with obvious modifications) for a $l \times m$ frequency table of l rows and m columns, where $2 \leqslant l \leqslant n$, $2 \leqslant m \leqslant n$.

3. KENDALL'S SCORE S

For a pair of rankings of n members, Kendall (1955) defined the *score* associated with any two of the n members as $+1$ if they are ranked in the same order by the two rankings, -1 if they are ranked in opposite order, and 0 if they are tied in either or both rankings. The total score S for the pair of rankings is the algebraic sum of the $\frac{1}{2}n(n-1)$ contributions obtained from all such pairs of members.

For the pair of rankings shown in Table 1, the score S is the algebraic sum of all second-order determinants in $\{a\}$, thus

$$S = (a_{11}a_{22} - a_{21}a_{12}) + (a_{11}a_{23} - a_{21}a_{13}) + \ldots + (a_{23}a_{34} - a_{33}a_{24}). \tag{3·1}$$

For example, there are $a_{11}a_{23}$ ways of selecting an untied pair of members from the two cells (row 1, column 1) and (row 2, column 3), and each such pair contributes $+1$ to S. Similarly, each of the $a_{21}a_{13}$ untied pairs which can be selected from (row 2, column 1) and (row 1, column 3) contributes -1 to S.

In practice, the sum (3·1) may be found rapidly by grouping either the scores associated with each element a_{ij} and elements in lines below it, or those associated with each element a_{ij} and elements in columns to the right of it (Kendall, 1955, § 3·16).

We observe that a row or column whose elements are all zero may be inserted in or deleted from Table 1 without affecting S.

4. Probabilities in random sampling

We now regard Table 1 as the result of drawing a random sample (subject to restrictions on the marginal totals $\{u\}$, $\{v\}$) of n members from a large parent population. We suppose for the moment that each member of this population, if drawn, would belong in a definite cell in Table 1. If this is not true of the whole population then we confine our attention to the subpopulation for which it is true. (This restriction will be removed later.)

In practice, the investigator may or may not choose one or both of the sets of marginal totals before drawing the sample. When we wish to find the sampling distribution of S for Table 1, the question arises whether the terms in $\{u\}$, $\{v\}$ should be regarded as fixed in size and order for all possible samples, as though the investigator had chosen them in advance, or whether they should be subject to random fluctuations from one sample to another.

In the case of 2×2 tables this question has given rise to some controversy (Fisher, 1956), but it seems to be generally agreed that it is permissible, though perhaps not ideal, to regard both $\{u\}$ and $\{v\}$ as fixed. In this paper we regard $\{u\}$ and $\{v\}$ as fixed during sampling for tables of any size, mainly because of the enormous increase in complexity which results if we try to allow for fluctuations.

We now set up the null hypothesis H_0, that the properties U, V are independent in the population. This means that the probability that a member of the population will belong in a specified column is not affected by foreknowledge of the row in which it lies.

Let α_i be the probability that a member of the population will belong in row i, and β_j the probability that it will belong in column j. Then if H_0 is true, the probability that the member will belong in the cell in row i, column j is $\alpha_i\beta_j$. We recall that the population has been so restricted that every sample can be displayed in a table having at most 3 rows and at most 4 columns. In a random sample of n members with *unrestricted* marginal totals (save for the restriction in the preceding sentence), the probability of obtaining the particular cell frequencies shown in Table 1 is given by the multinomial distribution

$$\frac{n!}{a_{11}!a_{12}!\ldots a_{34}!}(\alpha_1\beta_1)^{a_{11}}(\alpha_1\beta_2)^{a_{12}}\ldots(\alpha_3\beta_4)^{a_{34}},$$

which readily simplifies to

$$\frac{n!}{a_{11}!a_{12}!\ldots a_{34}!}\alpha_1^{u_1}\alpha_2^{u_2}\alpha_3^{u_3}\beta_1^{v_1}\beta_2^{v_2}\beta_3^{v_3}\beta_4^{v_4}. \tag{4·1}$$

Now in sampling with unrestricted row totals, the probability of obtaining the particular row totals u_1, u_2, u_3 in that order is

$$\frac{n!}{u_1!\,u_2!\,u_3!}\alpha_1^{u_1}\alpha_2^{u_2}\alpha_3^{u_3}. \tag{4·2}$$

Hence on dividing (4·1) by (4·2) we find that the probability of obtaining the cell frequencies in Table 1 in sampling with $\{u\}$ fixed is

$$\frac{u_1!\,u_2!\,u_3!}{a_{11}!\,a_{12}!\,\dots\,a_{34}!}\beta_1^{v_1}\beta_2^{v_2}\beta_3^{v_3}\beta_4^{v_4}. \tag{4·3}$$

Again, in sampling with unrestricted column totals, the probability of obtaining the particular column totals v_1, v_2, v_3, v_4 in that order is

$$\frac{n!}{v_1!\,v_2!\,v_3!\,v_4!}\beta_1^{v_1}\beta_2^{v_2}\beta_3^{v_3}\beta_4^{v_4}, \tag{4·4}$$

and because of H_0 this is true whether the row totals have been restricted or not. Hence on dividing (4·3) by (4·4) we find that the probability of obtaining the cell frequencies in Table 1 in sampling with both $\{u\}$ and $\{v\}$ fixed is

$$\frac{u_1!\,u_2!\,u_3!\,v_1!\,v_2!\,v_3!\,v_4!}{n!\,a_{11}!\,a_{12}!\,\dots\,a_{34}!}. \tag{4·5}$$

The probability of obtaining the score S in such a sample is therefore

$$p(S;\{u\},\{v\}) = \Sigma\frac{u_1!\,u_2!\,u_3!\,v_1!\,v_2!\,v_3!\,v_4!}{n!\,a_{11}!\,a_{12}!\,\dots\,a_{34}!}, \tag{4·6}$$

where the summation extends over all permissible arrays $\{a\}$ which give rise to the score S. Now if we multiply both sides of (4·6) by a suitable constant, we may replace the probabilities p by relative frequencies P which are all integers. Two obvious ways of choosing this constant yield relative frequencies

$$P_1(S;\{u\},\{v\}) = \frac{n!}{v_1!\,v_2!\,v_3!\,v_4!}p(S;\{u\},\{v\}) = \Sigma\frac{u_1!\,u_2!\,u_3!}{a_{11}!\,a_{12}!\,\dots\,a_{34}!}, \tag{4·7}$$

$$P_2(S;\{u\},\{v\}) = \frac{n!}{u_1!\,u_2!\,u_3!}p(S;\{u\},\{v\}) = \Sigma\frac{v_1!\,v_2!\,v_3!\,v_4!}{a_{11}!\,a_{12}!\,\dots\,a_{34}!}. \tag{4·8}$$

We note that $$\Sigma P_1 = n!/v_1!\,v_2!\,v_3!\,v_4!, \quad \Sigma P_2 = n!/u_1!\,u_2!\,u_3!. \tag{4·9}$$

If the smallest such constant is chosen, so that the relative frequencies are integers having no common factor, then we shall denote these relative frequencies by $P(S;\{u\},\{v\})$.

We observe that the score S, the probabilities p and the relative frequencies P_1, P_2, P are all unaffected by the insertion or removal of a row or column of zeros in Table 1, or any number of such rows or columns. This observation enables us to remove the restriction placed on the parent population at the beginning of this section. For suppose that the members of the population can be ranked in a table of (for example) 5 rows and 7 columns. Then Table 1 may be regarded as a 5×7 table whose row totals are one of the 10 sequences $(u_1, u_2, u_3, 0, 0)$, $(u_1, u_2, 0, u_3, 0)$, …, $(0, 0, u_1, u_2, u_3)$, and whose column totals are one of the 35 sequences $(v_1, v_2, v_3, v_4, 0, 0, 0)$, $(v_1, v_2, v_3\,0, v_4, 0, 0)$, …, $(0, 0, 0, v_1, v_2, v_3, v_4)$. Thus the set of all possible samples whose non-zero marginal totals have the fixed values $\{u\}$, $\{v\}$ of

Table 1, fall into 350 subsets according to the distribution of the zero marginal totals among the non-zero totals. If we confine our attention to the samples of a single preassigned subset, the values of S, p, P_1, P_2, P are as given above. But these values are independent of the choice of subset. Therefore the division into subsets is irrelevant, and our results are true for sampling in which the non-zero marginal totals are fixed, irrespective of the number or position of empty rows or columns. We remark that this reasoning is valid only if the null hypothesis is true for the whole population; it is not sufficient that the null hypothesis be true merely for those classes which are represented in one sample.

5. The significance of S

The exact probability distribution of S for random sampling with fixed $\{u\}$, $\{v\}$ is given by (4·6) when the null hypothesis is true. It has been shown by Kendall (1955) that the mean value of S in this distribution is zero, that its variance is given by

$$\operatorname{var} S = \frac{1}{18}\Big[f(n) - \sum_i f(u_i) - \sum_j f(v_j)\Big] + \frac{1}{9g(n)}\Big[\sum_i g(u_i)\Big]\Big[\sum_j g(v_j)\Big] + \frac{1}{2h(n)}\Big[\sum_i h(u_i)\Big]\Big[\sum_j h(v_j)\Big], \tag{5·1}$$

where

$$f(x) = x(x-1)(2x+5),$$
$$g(x) = x(x-1)(x-2),$$
$$h(x) = x(x-1);$$

and that, if the number of tied elements in each ranking is small compared with n, the distribution tends to normality as n increases indefinitely. It is also known that for 2×2 tables the distribution of S tends to normality when u_1, u_2, v_1, v_2 all tend to infinity. It seems plausible that the distribution should tend to normality in the general case, except possibly when one of the ratios u_i/n and/or one of v_j/n tends to unity; but no formal proof of this property is known.

Knowing the probability distribution of S, we can determine the level of significance of an observed score in the usual way. Thus if Table 1 yields a positive score S_0, we can compute the probability of obtaining a score $S \geqslant S_0$ in random sampling with fixed $\{u\}$, $\{v\}$, and we call this probability the (one-tailed) significance level of S_0. If this probability is small, for example, less than 0·05 or 0·01, we may reject H_0 at the 5 % level or the 1 % level, respectively, in favour of the alternative hypothesis that there is some degree of positive monotonic correlation between the two rankings, that is, that the members of the population tend to be ranked in approximately the same order by the two rankings. This procedure may be modified in the usual ways to deal with negative values of S_0 or with a two-tailed test.

This test must be distinguished from a form of the chi-square test in which the test statistic is

$$X^2 = \sum_{i,j} (a_{ij} - A_{ij})^2 / A_{ij},$$

where $A_{ij} = u_i v_j / n$. For a table of l rows and m columns, it is known that, when n is large and the null hypothesis is true, X^2 conforms approximately to the chi-square distribution with $(l-1)(m-1)$ degrees of freedom. In the special case of a 2×2 table, this test is exactly equivalent to the two-tailed normal approximation to the S test. But for a table with more than two rows or columns the tests are quite different, as the S test depends on the order of

the rows and columns, while the chi-square test is independent of the order. For example, the 2×3 table with $n = 30$ and cell frequencies

$$\begin{matrix} 7 & 1 & 7 \\ 3 & 9 & 3 \end{matrix}$$

yields a value of X^2 significant at the 1 % level ($X^2 = 9{\cdot}6$ on 2 D.F.), indicating that there is almost certainly some kind of association between the properties U, V; but the non-significant score $S = 0$ shows that there is no ground for belief that there is any degree of monotonic correlation. On the other hand, the 3×3 table with $n = 27$ and cell frequencies

$$\begin{matrix} 5 & 3 & 1 \\ 3 & 3 & 3 \\ 1 & 3 & 5 \end{matrix}$$

yields a non-significant value of X^2, but the score $S = 96$ is significant at the 5 % level (two-tailed test), the standard deviation of S being 42·4. The S test is clearly more powerful than the chi-square test in discriminating between H_0 and the alternative hypothesis of a tendency to monotonic correlation, while the chi-square test is appropriate for testing for association of an unspecified kind.

6. Calculation of the distribution of S

Except in the case of tables with only two rows (or columns), the task of enumerating separately all permissible sets of cell frequencies in the sum (4·6), and evaluating S and the contribution to p for each, is excessively tedious, even for small samples such as $n = 6$. But we now show how the distribution (4·6) for a 3×4 table of n members can be expressed as a linear combination of certain distributions for either $n - u_3$ or $n - v_4$ members.

Of the set of all permissible arrays $\{a\}$ in Table 1 with fixed $\{u\}$, $\{v\}$, consider the subset in which the elements of the fourth column have the fixed values a_1, a_2, a_3. If we delete this column, we obtain the set of all permissible arrays in a 3×3 table with the fixed marginal totals $u_1 - a_1, u_2 - a_2, u_3 - a_3$; v_1, v_2, v_3. Let S' denote the score for a typical 3×3 array and S the score for the corresponding 3×4 array. Then by considering the terms of (3·1) which occur in S but not in S', we obtain

$$S = S' + [(u_1 - a_1) + (u_2 - a_2)]\, a_3 + [(u_1 - a_1) - (u_3 - a_3)]\, a_2 - [(u_2 - a_2) + (u_3 - a_3)]\, a_1,$$

that is,

$$S = S' + (u_1 + u_2)\, a_3 + (u_1 - u_3)\, a_2 - (u_2 + u_3)\, a_1. \tag{6·1}$$

The contribution of this subset to the sum (4·8) is

$$\frac{v_4!}{a_1!\, a_2!\, a_3!} \Sigma' \frac{v_1!\, v_2!\, v_3!}{a_{11}!\, a_{12}! \ldots a_{33}!} \equiv \frac{v_4!}{a_1!\, a_2!\, a_3!} P_2(S'; \{u\}', \{v\}'),$$

where $\{u'\} \equiv (u_1 - a_1, u_2 - a_2, u_3 - a_3)$, $\{v\}' \equiv (v_1, v_2, v_3)$, S' is related to S by (6·1), and the summation Σ' extends over all permissible 3×3 arrays whose score is S'. Finally, on summing these contributions over all possible subsets, that is, all possible arrangements of the v_4 elements in the fourth column, we obtain

$$P_2(S; \{u\}, \{v\}) = \Sigma \frac{v_4!}{a_1!\, a_2!\, a_3!} P_2(S'; \{u\}', \{v\}'), \tag{6·2}$$

where the summation extends over all permissible columns (a_1, a_2, a_3) whose sum is v_4.

Similarly, by considering the 2×4 arrays obtained by deleting the third row of Table 1, it may be shown that

$$P_1(S; \{u\}, \{v\}) = \Sigma \frac{u_3!}{b_1! b_2! b_3! b_4!} P_1(S''; \{u\}'', \{v\}''), \tag{6·3}$$

where the summation extends over all permissible rows (b_1, b_2, b_3, b_4) whose sum is u_3, $\{u\}'' \equiv (u_1, u_2)$, $\{v\}'' \equiv (v_1 - b_1, v_2 - b_2, v_3 - b_3, v_4 - b_4)$, and S'' is related to S by

$$S = S'' + (v_1 + v_2 + v_3) b_4 + (v_1 + v_2 - v_4) b_3 + (v_1 - v_3 - v_4) b_2 - (v_2 + v_3 + v_4) b_1. \tag{6·4}$$

For a 'table' of only one row (or column), it is obvious that $S = 0$ is the only possible score, so that $p(S) = 1$ for $S = 0$ and $p(S) = 0$ for $S \neq 0$. Hence we may use (6·2) or (6·3) to find the distribution of S for tables of two rows (or columns), then for tables of three rows (or columns), and so on.

In this way the distributions of S for all possible marginal totals for $n \leqslant 6$ have been computed. They are given in Table 4. The values of $P(S)$ for each of the 241 distributions in this Table were checked by computing the moments of orders zero, one and two, the second moment being checked against the value given by (5·1). The extreme values of S and the corresponding values of P were also checked in each case by the method of § 8.

For $n > 6$, the distribution of S may be found by a single application of (6·2) or (6·3) with the aid of Table 4, provided that the number of members in the sample can be reduced to 6 or fewer by delecting the first or last row or column; otherwise two or more applications of (6·2) or (6·3) are necessary. However, for $n \geqslant 9$ the calculation is rarely necessary (see § 9).

7. Examples of the use of table 4

Example 7·1. Let 6 objects be ranked or graded according to the two properties U, V, with the result

A	A	B	C	C	D
a	a	c	b	e	d

(7·1)

where A, B, C, D denote grades of U and a, b, c, d, e denote grades of V, both in order of decreasing degree. (Note that the letters b, e in (7·1) could equally well have been written in reverse order, since the objects concerned are ranked equal according to U.) In the form of Table 1, we obtain the 4×5 table

2	0	0	0	0	2
0	0	1	0	0	1
0	1	0	0	1	2
0	0	0	1	0	1
2	1	1	1	1	6

We wish to test the null hypothesis that there is no association between U, V, against the alternative hypothesis that there is some degree of monotonic correlation between U, V.

The class frequencies are $\{u\} = (2, 1, 2, 1)$, $\{v\} = (2, 1, 1, 1, 1)$ and the score $S = +9$. Referring to the last page of Table 4, we locate the group of lines for which $\{u\} = (2, 1, 2, 1)$ (first column), and within this group we find the line for which $\{v\} = (2, 1, 1, 1, 1)$ (second column). Adding the values of $P(9)$, $P(10)$, $P(11)$ and $P(13)$ given on this line, and dividing by ΣP, we obtain $(3 + 2 + 2 + 1)/180 = 2/45 \simeq 0{\cdot}044$, so that the one-tailed significance level of the score $+9$ is 4·4 %. We conclude, with a degree of confidence measured by the

significance level of 4·4 %, that the null hypothesis of independence between U, V may be rejected in favour of the alternative hypothesis that there is some degree of positive monotonic correlation between U, V.

If the properties U, V had been taken in reverse order, the second line of (7·1) would have been written first, and we should have obtained $\{u\} = (2, 1, 1, 1, 1)$, $\{v\} = (2, 2, 2, 1)$, $S = +9$. The distribution $P(S)$ for these $\{u\}$, $\{v\}$ is not listed separately in Table 4, because it is identical with the distribution for $\{u\} = (2, 1, 2, 1)$, $\{v\} = (2, 1, 1, 1, 1)$.

If the order of *increasing* degree had been adopted as the natural order for both U and V, we should have obtained $\{u\} = (1, 2, 1, 2)$, $\{v\} = (1, 1, 1, 1, 2)$, $S = +9$. The distribution $P(S)$ for these $\{u\}$, $\{v\}$ is again identical with that for $\{u\} = (2, 1, 2, 1)$, $\{v\} = (2, 1, 1, 1, 1)$.

Example 7·2. Let $n = 5$, $\{u\} = (1, 1, 2, 1)$, $\{v\} = (1, 2, 1, 1)$, $S = +6$. Since the class frequencies $(1, 1, 2, 1)$ do not occur in Table 4 in that order, we must first reverse the order of the terms in $\{u\}$, and this involves a change in the sign of S. Hence we seek the level of significance of the score $S = -6$ for the distribution for which $\{u\} = \{v\} = (1, 2, 1, 1)$. Referring to the appropriate line on the first page of Table 4, we add $P(-6)$ and $P(-8)$ and divide the sum by ΣP, giving $(4+2)/60 = 1/10$. Therefore the one-tailed significance level of the observed score is 10 %.

Example 7·3. Let $n = 6$, $\{u\} = (1, 1, 2, 1, 1)$, $\{v\} = (1, 2, 3)$, $S = -8$. Before entering Table 4, we must interchange $\{u\}$ and $\{v\}$ and reverse the order of the terms in $(1, 2, 3)$. The second of these operations normally entails a change in the sign of S, but as the sequence $(1, 1, 2, 1, 1)$ is symmetrical, the distribution $P(S)$ is also symmetrical, and the sign of S is not relevant to the determination of the significance level. We now enter Table 4 with $\{u\} = (3, 2, 1)$, $\{v\} = (1, 1, 2, 1, 1)$, and $S = 8$ or -8, and obtain the one-tailed significance level $(2+2)/60 = 1/15$. The two-tailed significance level is 2/15.

Example 7·4. Let $n = 6$, $\{u\} = (1, 3, 2)$, $\{v\} = (1, 1, 1, 1, 1, 1)$, $S = 7$. Since the second ranking is untied, the order of the terms in $\{u\}$ is immaterial. The required distribution $P(S)$ is grouped in Table 4 with the first $\{u\}$ whose terms are a permutation of $(1, 3, 2)$, namely, $\{u\} = (3, 2, 1)$. We obtain for the one-tailed significance level $(4+2+1)/60 = 7/60 \simeq 0{\cdot}12$, in agreement with the value given by Sillitto (1947, Table 2). The two-tailed significance level is 7/30.

Example 7·5. Let $n = 7$, $\{u\} = (3, 1, 3)$, $\{v\} = (2, 1, 3, 1)$. It is required to find the distribution of S. The simplest method is to use the expansion (6·2) in terms of distributions of 6 members, obtained by deleting the last column of the frequency table of 7 members. Since $v_4 = 1$, the only possible sets of cell frequencies for the last column (a_1, a_2, a_3) are $(0, 0, 1)$, $(0, 1, 0)$ and $(1, 0, 0)$, and the corresponding row totals $\{u\}'$ for the reduced tables are $(3, 1, 2)$, $(3, 0, 3)$ and $(2, 1, 3)$. In each case $v_4!/a_1!\,a_2!\,a_3! = 1$. Hence (6·2) yields

$$P_2(S, 313, 2131) = P_2(S', 312, 213) + P_2(S'', 303, 213) + P_2(S''', 213, 213),$$

where $S = S' + 4$, $S = S'' + 0$, $S = S''' - 4$, by (6·1). The contribution $P_2(S', 312, 213)$ is found from the distribution P_2 for $\{u\} = (3, 1, 2)$, $\{v\} = (3, 1, 2)$ by changing the sign of each score (due to reversal of $\{v\}$) and then adding 4 to each score. The contribution $P_2(S'', 303, 213)$ is the same as the distribution P_2 for $\{u\} = (3, 3)$, $\{v\} = (3, 1, 2)$, since the empty row may be ignored. The contribution $P_2(S''', 213, 213)$ is found from the distribution P_2 for

$$\{u\} = \{v\} = (3, 1, 2)$$

by subtracting 4 from each score. The contributions are added as shown in Table 2. Some negative values of S have been omitted from Table 2 to conserve space, but the last line is easily completed by symmetry. The result may be checked by verifying that

$$\Sigma P_2 = n!/u_1!\,u_2!\,u_3! = 140 \text{ [see (4·9)]}, \quad \text{that} \quad \Sigma S^2 P_2 = 140\,(\text{var}\,S) = 4692 \text{ [see (5·1)]},$$

and that the extreme scores and their probabilities agree with the values found by the method of § 8 (see example 8·1). We conclude that the one-tailed significance levels of the scores 13, 11, 9, 8, ... are 3/140, 6/140, 7/140, 19/140, ...

Table 2. *Summation for example* 7·5

S	...	−4	−3	−2	−1	0	1	2	3	4	5	6	7	8	9	10	11	12	13
$P_2(S', 312, 213)$	...	2	3	.	.	12	.	6	.	6	12	.	.	12	.	.	3	.	3
$P_2(S'', 303, 213)$	...	.	.	6	.	.	.	6	.	.	3	.	.	.	1	.	.	.	.
$P_2(S''', 213, 213)$	...	6	.	6	.	12	.	.	3	2	.	.	1	.	.	.	.	.	.
$P_2(S, 313, 2131)$	...	8	3	12	.	24	.	12	3	8	15	.	1	12	1	.	3	.	3

Example 7·6. Let $n = 8$, $\{u\} = (1, 3, 2, 2)$, $\{v\} = (1, 2, 1, 4)$. It is required to find the distribution of S. We use (6·3) to express $P_1(S, 1322, 1214)$ in terms of the distributions obtained by deleting the last row. We first list the eight permissible ways of inerting the two members in the last row (see below). Beside each we write the corresponding term on the right of (6·3), complete with its multinomial coefficient. We then rewrite each of these terms, omitting empty columns (if any) and making any reversals or interchanges in $\{u\}$, $\{v\}$ needed to make immediate entry into Table 4 possible. (If $\{u\}$, $\{v\}$ are interchanged, P_1 must be changed to P_2, and if the sign of S' has to be changed, this must be noted.) Finally, we write the relation between S and S' in each case, giving the following scheme:

0002	$1P_1(S', 132, 1212)$	$1P_1(S', 231, 2121)$	$S = S' + 8$
0011	$2P_1(S', 132, 1203)$	$2P_2(S', 321, 231)$	$S = S' + 3$
0101	$2P_1(S', 132, 1113)$	$2P_1(S', 231, 3111)$	$S = S' + 0$
1001	$2P_1(S', 132, 0213)$	$2P_2(S', 312, 231)$	$S = S' - 3$
0110	$2P_1(S', 132, 1104)$	$2P_2(S', 411, 231)$	$S = S' - 5$
1010	$2P_1(S', 132, 0204)$	$2P_2(S', 42, 231)$	$S = S' - 8$
0200	$1P_1(S', 132, 1014)$	$1P_2(S', 411, 231)$	$S = S' - 8$
1100	$2P_1(S', 132, 0114)$	$2P_2(S', 411, 231)$	$S = S' - 11$

The increments to be added to S' to give S were found rapidly by writing down the 2×4 arrays

1212	1203	1113	0213	1104	0204	1014	0114
0002	0011	0101	1001	0110	1010	0200	1100

and finding the score, that is, the sum of the second-order determinants, in each. This is more convenient than direct substitution in (6·4). The eight distributions are now summed as shown in Table 3 (in which some values of S have been omitted to conserve space). Each of these eight distributions is derived from the corresponding distribution $P(S)$ in Table 4 by adding the appropriate increment to each score and multiplying each P by the multinomial coefficient and by the factor P_1/P or P_2/P. The sum of these eight distributions is the required distribution $P_1(S, 1322, 1214)$. The result may be checked by verifying that

$\Sigma P_1 = n!/v_1!\,v_2!\,v_3!\,v_4! = 840$ [see (4·9)], that $\Sigma S P_1 = 0$, that $\Sigma S^2 P_1 = 840\,(\mathrm{var}\,S) = 43{,}090$ [see (5·1)], and that the extreme scores and their probabilities agree with those found by the method of § 8. Unfortunately none of these checks would reveal a mistake caused by the erroneous use of a distribution with (for example) $\{u\} = (321)$ instead of (312) in constructing Table 3, or omitting to change the sign of S' in such a distribution, unless the mistake affected one of the extreme scores. Therefore the computer should take great care to avoid such mistakes. The only certain safeguard is to repeat the entire calculation by another method, that is, by deleting a different row or column.

Table 3. *Summation for example* 7·6

− 19	− 18	− 17	− 16	− 15	− 14	− 13	− 12	− 11	− 10	− 9	...	+ 9	+ 10	+ 11	+ 12	+ 13	+ 14	+ 15	+ 16	+ 17	+ 18	+ 19	Σ
.	.	.	.	.	.	.	.	.	.	.	...	24	12	12	.	27	.	.	12	.	.	3	180
.	.	.	.	.	.	.	.	.	.	.	...	12	.	.	6	.	.	.	.	.	.	.	120
.	.	.	.	.	.	.	.	.	.	12	...	12	.	.	.	.	.	.	.	.	.	.	240
.	.	.	.	.	.	.	6	.	.	12	...	.	.	.	.	.	.	.	.	.	.	.	120
.	.	.	.	.	.	4	.	.	.	12	...	.	.	.	.	.	.	.	.	.	.	.	60
.	.	.	2	.	.	.	.	12	.	.	...	.	.	.	.	.	.	.	.	.	.	.	30
.	.	.	2	.	.	.	6	.	6	.	...	.	.	.	.	.	.	.	.	.	.	.	30
4	.	.	.	12	.	12	.	4	.	16	...	.	.	.	.	.	.	.	.	.	.	.	60
4	.	.	4	12	.	16	12	16	6	52	...	48	12	12	6	27	.	.	12	.	.	3	840
0·48	.	.	0·95	2·4	.	4·3	5·7	7·6	8·3	15	.	14	8·6	7·1	5·7	5·0	.	.	1·8	.	.	0·36	L (%)
0·49	.	.	1·5	2·1	.	4·0	5·4	7·1	9·2	12	.	12	9·2	7·1	5·4	4·0	.	.	1·5	.	.	0·49	L' (%)

The whole calculation described above, including all checks, was completed in about 1 hr.

In the last two lines of Table 3 are given the one-tailed significance levels L computed from $P_1(S)$, and the corresponding approximate levels L' computed from the normal approximation to $p(S)$, with $\mathrm{var}\,S = 51{\cdot}3$ and a continuity correction of $\frac{1}{2}$ subtracted from each positive score and added to each negative score. Both L and L' are expressed as percentages correct to two significant figures. It is seen that, except for two of the four extreme scores (which can be treated separately by the method of § 8), the differences between L and L' never exceed one-fifth of L.

8. Significance levels of extreme scores

In the distribution $p(S, \{u\}, \{v\})$, let S_1 denote the highest positive score which occurs, S_2 the highest but one, and S_3 the highest but two. It will now be shown how we can always find fairly quickly the exact significance levels of S_1 and S_2 (and sometimes of S_3), and of the corresponding negative scores, working from first principles.

We first consider the effect on the score of a transposition of two of the n members in one ranking. Let a set of four cell frequencies belonging to two adjacent rows and two adjacent columns in Table 1 be

$$\begin{matrix} a & b \\ c & d. \end{matrix}$$

If a, d are both positive (i.e. not zero), we may move one of the a members down one row and add it to c, provided that we simultaneously move one of the d members up one row and

add it to b, so that the marginal totals remain unchanged. This transposition may be denoted by

$$\left.\begin{matrix} a & b & a-1 & b+1 \\ c & d & c+1 & d-1 \end{matrix}\right\rangle \qquad (8{\cdot}1)$$

and it is easy to show that its effect on the total score S is to reduce S by $a+b+c+d$ units, provided that all the other cell frequencies remain unchanged. When the rows or columns affected by the transposition are not adjacent, we find similarly that the transposition

$$\left.\begin{matrix} a & e_1 & e_2 & b & a-1 & e_1 & e_2 & b+1 \\ c & e_3 & e_4 & d & c+1 & e_3 & e_4 & d-1 \end{matrix}\right\rangle \qquad (8{\cdot}2)$$

reduces S by $(a+b+c+d)+2(e_1+e_2+e_3+e_4)$, and more generally the transposition

$$\left.\begin{matrix} a & e_1 & e_2 & b & a-1 & e_1 & e_2 & b+1 \\ f_1 & x_1 & x_2 & f_2 & f_1 & x_1 & x_2 & f_2 \\ c & e_3 & e_4 & d & c+1 & e_3 & e_4 & d-1 \end{matrix}\right\rangle \qquad (8{\cdot}3)$$

reduces S by $(a+b+c+d)+2(e_1+e_2+e_3+e_4+f_1+f_2)+4(x_1+x_2)$, and similarly for any other numbers of intervening rows and columns.

To find the extreme score S_1 for Table 1, we choose the cell frequencies $\{a\}$ in such a way that the n members are ranked in the same order by the two rankings, except in so far as this is prevented by ties. This means that no two cells with positive cell frequencies can be joined by a line sloping upwards from left to right, and that there are no negative terms in (3·1). Having written down this array $\{a\}$ (which is clearly unique), we at once obtain S_1 from (3·1) and $p(S_1)$ from (4·6), the sum on the right of (4·6) now comprising only one term.

To find the score S_2, we examine all possible transpositions of the form (8·1) in the array $\{a\}$ whose score is S_1. Let m be the smallest value of $a+b+c+d$ which occurs (for a, b, c, d in adjacent rows and columns and a, d both positive). (Note that in the present case at least one of b, c must be zero.) Then $S_2 = S_1 - m$, and $p(S_2)$ is given by (4·6), the number of terms on the right of (4·6) now being equal to the number of distinct transpositions which reduce the score by m. The contribution to $p(S_2)$ due to each of these terms can be written down by inspection as a simple rational multiple of $p(S_1)$, in the form

$$\frac{ad}{(b+1)(c+1)}p(S_1).$$

In practice it is easier to work in terms of $P_1(S)$ or $P_2(S)$, whichever is smaller.

The determination of all arrays $\{a\}$ whose score is S_3 may involve transpositions of the form (8·2) or (8·3) or one or two transpositions of the form (8·1), and is often difficult, but in some cases it is quite simple, as in the following example.

Example 8·1. Let $n = 7$, $\{u\} = (3, 1, 3)$, $\{v\} = (2, 1, 3, 1)$. Here $\Sigma P_1(S) = 7!/2!\,3! = 420$, while $\Sigma P_2(S) = 7!/3!\,3! = 140$, so we choose to work in terms of $P_2(S)$. In the array whose score is S_1 (see below), the transposition

$$\begin{matrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \end{matrix}$$

reduces the score by two units, and the transposition

$$\begin{matrix} 1 & 0 & 0 & 1 \\ 2 & 1 & 3 & 0 \end{matrix}$$

reduces the score by four units. Thus we obtain the arrays

2	1	0	0	3		2	0	1	0	3		2	1	0	0	3
0	0	1	0	1		0	1	0	0	1		0	0	0	1	1
0	0	2	1	3		0	0	2	1	3		0	0	3	0	3
2	1	3	1	7		2	1	3	1	7		2	1	3	1	7

with scores 13, 11, 9; and we see that these are the only arrays giving scores $\geqslant 9$. (For example, the transpotition

$$\begin{matrix} 2 & 1 & 0 & & 1 & 1 & 1 \\ 0 & 0 & 1 & \rightarrow & 1 & 0 & 0 \end{matrix}$$

in the first array reduces the score by 5.) Therefore

$$S_1 = 13, \quad S_2 = 11, \quad S_3 = 9, \quad P_2(S_1) = 2!\,3!/2!\,2! = 3, \quad P_2(S_2) = 3, \quad P_2(S_3) = 1,$$

and the corresponding one-tailed significance levels are 3/140, 6/140, 7/140, in agreement with the results obtained by a longer method in example 7·5.

Example 8·2. Let $n = 10$, $\{u\} = (3, 4, 3)$, $\{v\} = (2, 3, 3, 2)$. Here

$$\Sigma P_2(S) = 10!/3!\,4!\,3! = 4200.$$

In the array whose score is S_1 (see below), there are precisely four transpositions which reduce the score by 5. Thus we obtain the arrays

2	1	0	0	1	2	0	0	2	0	1	0	2	1	0	0	2	1	0	0
0	2	2	0	1	1	2	0	0	3	1	0	0	1	3	0	0	2	1	1
0	0	1	2	0	0	1	2	0	0	1	2	0	1	0	2	0	0	2	1

with scores $S_1 = 29$, $S_2 = 24$ (four times). Therefore

$$P_2(S_1) = 2!\,3!\,3!\,2!/(2!)^4 = 9, \quad P_2(S_2) = 18+6+6+18 = 48,$$

and the corresponding one-tailed significance levels are 9/4200 and 57/4200. Perhaps less obviously, there are precisely four ways of obtaining the score 20 (in each case by a further transposition), from which we can show that $S_3 = 20$, $P_2(S_3) = 6+6+6+6 = 24$, and the significance level of S_3 is 81/4200.

9. THE NORMAL APPROXIMATION

In example 7·6 and Table 3 we have already indicated the accuracy of the normal approximation in a certain distribution of 8 members with irregularly tied rankings. Examination of several other distributions of 9, 10 and 12 members has indicated that for a given number of rows and columns the accuracy improves with increasing n, and that for a given value of n, the accuracy improves with increasing numbers of rows or columns. Since the interval between scores (other than the extreme scores) is usually one unit, it is advisable to apply a continuity correction of half a unit except in the special cases described below.

Example 9·1. Let $n = 10$ and $\{u\} = \{v\} = (2, 2, 2, 2, 2)$. The exact distribution $p(S)$ was computed with the aid of Table 4 by a two-stage application of (6·2). The corresponding one-tailed significance levels L, and the approximate significance levels L' given by the

normal approximation with a continuity correction of $\frac{1}{2}$ and $\operatorname{var} S = 115{\cdot}6$, are shown below for all scores for which L' lies between 0·1 and 10 %:

S	15	16	17	18	19	20	21	22	23	
L (%)	9·6	7·3	6·7	5·0	4·2	3·7	2·6	2·1	1·78	
L' (%)	8·9	7·5	6·3	5·1	4·3	3·5	2·8	2·3	1·82	
S	24	25	26	27	28	29	30	31	32	33
L (%)	1·19	0·99	0·65	0·54	0·40	0·21	0·15	0·12	0·08	0·04
L' (%)	1·44	1·13	0·89	0·69	0·53	0·40	0·31	0·23	0·17	0·13

The approximation when $L > 1$ % is evidently so good that the computation of the exact distribution is scarcely necessary. When $L < 1$ % the errors in L' become large, but in this region the research worker is not usually very much concerned with the exact significance level; further, because L' is too large, the errors are 'safe' in the sense that they will not mislead the research worker into attributing significance to non-significant scores.

Example 9·2. Let $n = 10$ and let both rankings be untied. The exact significance levels L (from Kendall, 1955, Appendix Table 1), and the approximate significance levels L' given by the normal approximation with a continuity correction of 1 and $\operatorname{var} S = 125$, are shown below for all scores for which L' lies between 0·1 and 10 %:

S	17	19	21	23	25	27	29	31	33	35
L (%)	7·8	5·4	3·6	2·3	1·4	0·83	0·46	0·23	0·11	0·05
L' (%)	7·6	5·4	3·7	2·5	1·6	1·00	0·61	0·37	0·21	0·12

Comparison of the errors in L' in examples 9·1 and 9·2 shows that the presence of the ties in the rankings in example 9·1 does not greatly affect the accuracy of the normal approximation.

The appropriate correction for continuity is usually 1/2, but the following exceptional cases should be noted.

If one ranking (say the second) is untied, then it follows from the theory of § 8 that every transposition alters S by an even number, since every column in $\{a\}$ contains precisely one member. Hence the scores occur at intervals of 2 units, and the correction for continuity is unity. This result also holds when both rankings are untied.

If one ranking (say the first) is a dichotomy, then any transposition will alter S by one of the numbers v_1+v_2, v_2+v_3, v_3+v_4, ... or by the sum of two or more of these numbers. Hence if the highest common factor of these numbers is h, any two successive values of S will differ by h (or some whole multiple of h), and the correction for continuity is $\frac{1}{2}h$. For example, if $\{v\} = (3, 1, 3, 1)$, $h = 4$; if $\{v\} = (3, 1, 1, 3)$ or $(1, 3, 3, 1)$, $h = 2$; if $\{v\} = (3, 2, 3, 7)$, $h = 5$; if $\{v\} = (1, 1, \ldots, 1)$, $h = 2$; if $\{v\} = (t, t, \ldots, t)$, $h = 2t$; if $\{v\} = (v_1, v_2)$, $h = v_1 + v_2 = n$. This result unifies and generalizes three results given by Kendall (1955, § 4·12 (*a*), (*b*), (*c*)).

Example 9·3. Let $n = 12$, $\{u\} = (7, 5)$, $\{v\} = (6, 3, 3)$. Here $\operatorname{var} S = 127{\cdot}6$, $h = 3$, and the correction for continuity is 3/2. In the following results L, L' have the same meanings as before and some of the non-significant scores are omitted:

S	−30	−21	−15	−12	...	+12	+15	+18	+24	+27	+33
L (%)	0·76	6·4	12·1	19·7	...	14·1	11·8	9·8	3·0	0·8	0·38
L' (%)	0·58	4·2	11·6	17·6	...	17·6	11·6	7·2	2·3	1·2	0·27

It is seen that the normal approximation is not very good. However, the exact values of L for the six extreme scores $S \leqslant -15$ and $S \geqslant +24$ can all be found without difficulty by the method of § 8, and since these include all the significant scores, the normal approximation need not be used.

Several other distributions with $n = 9$, 10 and 12 were examined in this way. In some cases it was found that the normal approximation was not required, for the same reason as in example 9·3. In the remaining cases it was found that the normal approximation (with appropriate correction for continuity) was reasonably good at least in the region $L \geqslant 1\ \%$, as in examples 9·1 and 9·2. We conclude that for $n \geqslant 9$ it is seldom necessary to compute the exact distribution of S by the method of examples 7·5 and 7·6, except when great accuracy is required.

Finally, we comment further on the correction for continuity when the first ranking is a dichotomy and the second is tied in any manner. Whitfield (1947) suggests that the mean interval between successive values of S is approximately the mean of the numbers v_1+v_2, v_2+v_3, v_3+v_4, ..., so that the continuity correction is one-half of this mean interval. But the mean interval computed in this way is usually too large, as it gives the mean change in S produced by a *single* transposition involving adjacent columns, whereas in most cases a change from any value of S to the next attainable value involves two or more simultaneous transpositions which change S by a number of the form

$$(v_1+v_2)-(v_2+v_3) \quad \text{or} \quad (v_1+v_2)-(v_3+v_4).$$

Thus in example 9·3, the true mean interval between successive values of S (omitting the extreme values of $+33$ and -30, for which the continuity correction is never required) was found to be 4, which is closer to the highest common factor of $6+3$ and $3+3$ ($h = 3$) than to their mean ($= 7{\cdot}5$).

Example 9·4. Whitfield (1947) and Kendall (1955, examples 3·4 and 4·5) discuss a sample with $n = 17$ whose rankings may be displayed in the following 2×7 table:

4	0	2	4	2	0	0	12
0	1	0	1	1	1	1	5
4	1	2	5	3	1	1	17

Here $S = 34$, $\operatorname{var} S = 344{\cdot}6$, $h = 1$, and the mean of $4+1, 1+2, \ldots, 1+1$ is 29/6. With a continuity correction of 1/2 we obtain $L' = 3{\cdot}54\ \%$, but with a correction of 29/12, $L' = 4{\cdot}45\ \%$ (both one-tailed). The exact distribution $P_2(S)$ for $S \geqslant 23$ was found from (4·8) by enumerating the possible ways of arranging the five members in the second row, and yielded the exact significance level $L = 3{\cdot}36\ \%$ for $S = 34$. (Whitfield obtained $L = 3{\cdot}68\ \%$.) Comparison of L with the two values of L' shows that the continuity correction of 1/2 is the more appropriate. The reason for this is that every integral score in (at least) the range $23 \leqslant S \leqslant 42$ is attained, with the fortuitous exception of $S = 35$.

10. Relations with other tests

In the case of a 2×2 table only one kind of association is possible, since the cell frequencies have only one degree of freedom. The S test is then equivalent to Fisher's exact probability test, and the normal approximation to the S test is equivalent to the chi-square test.

In the case of a $2 \times m$ table, where the first ranking is a dichotomy and the second may or may not contain ties, the S test is equivalent to the Mann–Whitney U test, in which U is the sum of the positive terms only (or negative terms only) in the sum of the form (3·1) defining S. For example, if $n = 6$ and $\{u\} = (3, 3)$ and $\{v\} = (1, 1, 1, 1, 1, 1)$, the significance levels of $S = 9, 7, 5, 3$ and 1 are identical with the significance levels of $U = 0, 1, 2, 3$ and 4, respectively.

Hence the S test includes as special cases the Fisher test for association in 2×2 contingency tables, and the Mann–Whitney U test. For descriptions of these tests, bibliography and tables of distributions and significance points, see Siegel (1956).

References

Fisher, R. A. (1956). *Statistical Methods and Scientific Inference*. Edinburgh: Oliver and Boyd.

Kendall, M. G. (1955). *Rank Correlation Methods* (2nd edition). London: C. Griffin and Co. Ltd.

Siegel, S. (1956). *Nonparametric Statistics*. New York: McGraw-Hill.

Sillitto, G. P. (1947). The distribution of Kendall's τ coefficient of rank correlation in rankings containing ties. *Biometrika*, **34**, 36–40.

Whitfield, J. W. (1947). Rank correlation between two variables, one of which is ranked, the other dichotomous. *Biometrika*, **34**, 292–6.

Table 4. *Distribution of Kendall's score (of rank sum) S, for samples of 2, 3, 4, 5 and 6 members*

Explanation of the table

Each of the 241 distributions occupies one line of the table. The distributions for samples of $n = 2, 3, 4$ and 5 members are shown on the first page, and the distributions for samples of $n = 6$ members are shown on the remaining three pages. For each distribution, the figures in the columns from left to right show:

$\{u\}$, the class frequencies in the first ranking (or row totals).

$\{v\}$, the class frequencies in the second ranking (or column totals).

$P(S)$, the relative frequency of obtaining the score S (shown at the head of the column), in random sampling with fixed $\{u\}$, $\{v\}$ from a population in which the criteria of ranking are independent.

ΣP, the sum of the relative frequencies. Division by ΣP converts the relative frequencies $P(S)$ to probabilities.

P_1/P, the factor by which $P(S)$ must be multiplied to obtain $P_1(S)$, the relative frequencies such that $\Sigma P_1 = n!/(v_1!\, v_2! \ldots)$. Similarly $\Sigma P_2 = n!/(u_1!\, u_2! \ldots)$. The numbers P_1, P_2 are only required when computing a distribution for $n > 6$.

If the observed sets of class frequencies $\{u\}$, $\{v\}$ do not occur in Table 4, they must be modified in accord with one or more of the following rules before entering the Table:

(*a*) If the order of the terms in either $\{u\}$ or $\{v\}$ (but not both) is reversed, the sign of every score S is to be changed.

(*b*) If the order of the terms in both $\{u\}$ and $\{v\}$ is reversed, the distribution of S remains unchanged.

(*c*) If $\{u\}$ and $\{v\}$ are interchanged, then P_1 and P_2 are to be interchanged, but $P(S)$ remains unchanged.

(*d*) If one ranking contains no ties (i.e. every class frequency is unity), the distribution is the same for every permutation of the class frequencies in the other ranking.

(*e*) Empty classes in $\{u\}$ or $\{v\}$ or both may be ignored.

Test of significance for $n \leqslant 6$. To find the level of significance of an observed score, first locate the appropriate line of the table and denote the observed score (after changing its sign if necessary) by S_0. Add the tabulated relative frequencies P for the score S_0 and all scores to the right of S_0 (if S_0 is positive) or to the left of S_0 (if S_0 is negative). Divide the sum by the tabulated value of ΣP. The quotient is the 'one-tailed' significance level. If the distribution is symmetrical (which occurs when $\{u\}$ or $\{v\}$ or both are symmetrical) the quotient may be doubled to obtain the 'two-tailed' significance level. (See examples 7·1–7·4.)

Distributions for $n > 6$. For the sake of definiteness suppose that

$$\{u\} = (u_1, u_2, u_3) \quad \text{and} \quad \{v\} = (v_1, v_2, v_3, v_4), \quad \text{where} \quad \Sigma u = \Sigma v = n > 6.$$

Then in terms of distributions of $n - v_4$ members,

$$P_2(S, \{u\}, \{v\}) = \Sigma \frac{v_4!}{a_1!\, a_2!\, a_3!} P_2(S', \{u\}', \{v\}'),$$

where $\{u\}' = (u_1 - a_1, u_2 - a_2, u_3 - a_3)$, $\{v\}' = (v_1, v_2, v_3)$, and the score S is obtained from S' by adding the scores associated with the cell frequencies a_1, a_2, a_3 in the last column, thus

$$S = S' + (u_1 + u_2)\, a_3 + (u_1 - u_3)\, a_2 + (-u_2 - u_3)\, a_1;$$

and the summation extends over all permissible sets of non-negative integers a_1, a_2, a_3 whose sum is v_4.

Similarly in terms of distributions of $n - u_3$ members

$$P_1(S, \{u\}, \{v\}) = \Sigma \frac{u_3!}{b_1!\, b_2!\, b_3!\, b_4!} P_1(S', \{u\}', \{v\}'),$$

where $$\{u\}' = (u_1, u_2), \quad \{v\}' = (v_1 - b_1, \ldots, v_4 - b_4),$$

and $$S = S' + (v_1 + v_2 + v_3)\, b_4 + (v_1 + v_2 - v_4)\, b_3 + (v_1 - v_3 - v_4)\, b_2 + (-v_2 - v_3 - v_4)\, b_1.$$

(See examples 7·5 and 7·6.)

Table 4

See explanation on opposite page 166

		$P(S)$																					
		$-$	$-$	$-$	$-$	$-$	$-$	$-$	$-$	$-$		$+$	$+$	$+$	$+$	$+$	$+$	$+$	$+$	$+$			
$\{u\}$	$\{v\}$	9	8	7	6	5	4	3	2	1	0	1	2	3	4	5	6	7	8	9	ΣP	$\frac{P_1}{P}$	$\frac{P_2}{P}$
11	11	.	.	.	.	.	.	.	.	1	.	1	.	.	.	.	.	.	.	.	2	1	1
21	21	.	.	.	.	.	.	.	.	2	.	.	1	.	.	.	.	.	.	.	3	1	1
	111	.	.	.	.	.	.	.	1	.	1	.	1	.	.	.	.	.	.	.	3	2	1
111	111	.	.	.	.	.	.	1	.	2	.	2	.	1	.	.	.	.	.	.	6	1	1
31	31	.	.	.	.	.	.	.	.	3	.	.	.	1	.	.	.	.	.	.	4	1	1
	22	.	.	.	.	.	.	.	1	.	.	.	1	.	.	.	.	.	.	.	2	3	2
	211	.	.	.	.	.	.	.	2	.	.	1	.	1	.	.	.	.	.	.	4	3	1
	121	.	.	.	.	.	.	1	.	.	2	.	.	1	.	.	.	.	.	.	4	3	1
	1111	.	.	.	.	.	.	1	.	1	.	1	.	1	.	.	.	.	.	.	4	6	1
22	22	.	.	.	.	.	1	.	.	.	4	.	.	.	1	.	.	.	.	.	6	1	1
	211	.	.	.	.	.	1	.	.	2	.	2	.	.	1	.	.	.	.	.	6	2	1
	121	.	.	.	.	.	.	1	.	.	1	.	.	1	.	.	.	.	.	.	3	4	2
	1111	.	.	.	.	.	1	.	1	.	2	.	1	.	1	.	.	.	.	.	6	4	1
211	211	.	.	.	.	.	2	.	2	.	4	.	2	1	.	1	.	.	.	.	12	1	1
	121	.	.	.	.	.	2	.	2	1	2	1	2	.	2	.	.	.	.	.	12	1	1
	1111	.	.	.	.	1	.	2	.	3	.	3	.	2	.	1	.	.	.	.	12	2	1
121	121	.	.	.	.	1	.	.	4	.	2	.	4	.	.	1	.	.	.	.	12	1	1
1111	1111	.	.	.	1	.	3	.	5	.	6	.	5	.	3	.	1	.	.	.	24	1	1
41	41	.	.	.	.	.	.	.	.	4	.	.	.	.	1	.	.	.	.	.	5	1	1
	32	.	.	.	.	.	.	.	3	.	.	.	.	2	.	.	.	.	.	.	5	2	1
	311	.	.	.	.	.	.	.	3	.	.	.	1	.	1	.	.	.	.	.	5	4	1
	131	.	.	.	.	.	1	.	.	.	3	.	.	.	1	.	.	.	.	.	5	4	1
	221	.	.	.	.	.	.	2	.	.	.	2	.	.	.	1	.	.	.	.	.	5	6	1
	212	.	.	.	.	.	.	2	.	.	1	.	.	2	.	.	.	.	.	.	5	6	1
	2111	.	.	.	.	.	.	2	.	.	1	.	1	.	1	.	.	.	.	.	5	12	1
	1211	.	.	.	.	.	1	.	.	2	.	.	1	.	1	.	.	.	.	.	5	12	1
	11111	.	.	.	.	.	1	.	1	.	1	.	1	.	1	.	.	.	.	.	5	24	1
32	32	.	.	.	.	.	3	.	.	.	.	6	.	.	.	.	1	.	.	.	10	1	1
	311	.	.	.	.	.	3	.	.	.	3	.	3	.	.	.	1	.	.	.	10	2	1
	131	.	.	.	.	.	3	.	.	.	4	.	.	.	3	.	.	.	.	.	10	2	1
	221	.	.	.	1	.	.	.	4	.	.	2	1	.	.	2	.	.	.	.	10	3	1
	212	.	.	.	1	.	.	2	.	.	4	.	.	2	.	.	1	.	.	.	10	3	1
	2111	.	.	.	1	.	.	2	.	2	.	2	1	.	1	.	1	.	.	.	10	6	1
	1211	.	.	.	.	2	.	.	2	.	1	2	.	2	.	.	1	.	.	.	10	6	1
	11111	.	.	.	1	.	1	.	2	.	2	.	2	.	1	.	1	.	.	.	10	12	1
311	311	.	.	.	.	.	6	.	.	3	.	6	.	3	.	1	.	1	.	.	20	1	1
	131	.	.	.	.	3	.	3	.	1	6	1	.	3	.	3	.	.	.	.	20	1	1
	221	.	.	.	1	.	.	2	.	2	1	.	2	.	1	.	1	.	.	.	10	3	2
	212	.	.	.	1	.	1	.	1	2	.	2	1	.	1	.	1	.	.	.	10	3	2
	2111	.	.	.	2	.	2	.	4	.	4	1	2	2	.	2	.	1	.	.	20	3	1
	1211	.	.	.	2	.	2	1	2	2	2	1	4	.	2	1	.	1	.	.	20	3	1
	11111	.	.	1	.	2	.	3	.	4	.	4	.	3	.	2	.	1	.	.	20	6	1
131	131	.	.	1	.	.	.	6	.	.	6	.	.	6	.	.	.	1	.	.	20	1	1
	221	.	.	.	1	.	.	2	1	.	2	.	1	2	.	.	1	.	.	.	10	3	2
	212	.	.	.	.	1	.	.	1	.	1	.	1	.	.	1	.	.	.	.	5	6	4
	2111	.	.	.	2	.	2	1	2	2	2	2	2	1	2	.	2	.	.	.	20	3	1
	1211	.	.	1	.	1	2	.	4	1	2	1	4	.	2	1	.	1	.	.	20	3	1
221	221	.	.	.	4	.	.	6	.	.	9	.	4	.	4	2	.	.	1	.	30	1	1
	212	.	.	2	.	.	5	.	4	2	4	2	4	.	5	.	.	2	.	.	30	1	1
	2111	.	.	2	.	2	1	4	1	4	1	6	.	4	1	2	1	.	1	.	30	2	1
	1211	.	1	.	.	4	.	4	.	6	1	4	2	2	2	2	.	2	.	.	30	2	1
	11111	.	1	.	2	.	4	.	5	.	6	.	5	.	4	.	2	.	1	.	30	4	1
212	212	.	1	.	.	4	.	.	8	.	4	.	8	.	.	4	.	.	1	.	30	1	1
	2111	.	1	.	1	2	1	4	.	6	.	6	.	4	1	2	1	.	1	.	30	2	1
	1211	.	.	2	.	2	1	4	1	4	2	4	1	4	1	2	.	2	.	.	30	2	1
2111	2111	.	2	.	4	.	8	.	10	.	12	.	10	1	6	2	2	2	.	1	60	1	1
	1211	.	2	.	4	.	8	.	10	1	10	2	8	2	6	1	4	.	2	.	60	1	1
	11111	1	.	3	.	6	.	9	.	11	.	11	.	9	.	6	.	3	.	1	60	2	1
1211	1211	.	2	.	4	1	6	2	8	1	12	.	10	.	8	.	4	1	.	1	60	1	1
11111	11111	*	4	.	9	.	15	.	20	.	22	.	20	.	15	.	9	.	4	*	120	1	1

* For $\{u\} = \{v\} = (1, 1, 1, 1, 1)$, $P(-10) = 1$ and $P(+10) = 1$.

Table 4 (*cont.*)

See explanation on page 166

{u}	{v}	$P(S)$																	ΣP	$\frac{P_1}{P}$	$\frac{P_2}{P}$	{u}
		−9	−8	−7	−6	−5	−4	−3	−2	−1	0	+1	+2	+3	+4	+5	+6	+7	+8	+9				
51	51	.	.	.	.	.	.	.	.	5	.	.	.	.	.	1	.	.	.	.	6	1	1	
	42	.	.	.	.	.	.	.	2	.	.	.	.	.	1	.	.	.	.	.	3	5	2	41
	33	.	.	.	.	.	.	1	.	.	.	.	.	1	.	.	.	.	.	.	2	10	3	
	411	.	.	.	.	.	.	.	4	.	.	.	.	1	.	1	.	.	.	.	6	5	1	
	141	.	.	.	.	1	.	.	.	.	4	.	.	.	.	1	.	.	.	.	6	5	1	
	321	.	.	.	.	.	.	3	.	.	.	.	2	.	.	1	.	.	.	.	6	10	1	
	312	.	.	.	.	.	.	3	.	.	.	1	.	.	2	.	.	.	.	.	6	10	1	
	231	.	.	.	.	.	2	.	.	.	.	3	.	.	.	1	.	.	.	.	6	10	1	
	222	.	.	.	.	.	1	.	.	1	.	.	.	.	1	.	.	.	.	.	3	30	2	
	3111	.	.	.	.	.	.	3	.	.	.	1	.	1	.	1	.	.	.	.	6	20	1	
	1311	.	.	.	.	1	.	.	.	3	.	.	.	1	.	1	.	.	.	.	6	20	1	
	2211	.	.	.	.	.	2	.	.	.	2	.	.	1	.	1	.	.	.	.	6	30	1	14
	2121	.	.	.	.	.	2	.	.	1	.	.	2	.	.	1	.	.	.	.	6	30	1	
	2112	.	.	.	.	.	2	.	.	1	.	1	.	.	2	.	.	.	.	.	6	30	1	
	1221	.	.	.	.	1	.	.	2	.	.	.	2	.	.	1	.	.	.	.	6	30	1	
	21111	.	.	.	.	.	2	.	.	1	.	1	.	1	.	1	.	.	.	.	6	60	1	
	12111	.	.	.	.	1	.	.	2	.	.	1	.	1	.	1	.	.	.	.	6	60	1	
	11211	.	.	.	.	1	.	1	.	.	2	.	.	1	.	1	.	.	.	.	6	60	1	
	111111	.	.	.	.	1	.	1	.	1	.	1	.	1	.	1	.	.	.	.	6	120	1	
42	42	.	.	.	.	.	6	.	.	.	.	.	8	.	.	.	.	.	1	.	15	1	1	
	33	.	.	.	1	.	.	.	.	.	3	.	.	.	.	.	1	.	.	.	5	4	3	
	411	.	.	.	.	.	6	.	.	.	.	4	.	4	.	.	.	.	1	.	15	2	1	
	141	.	.	.	.	4	.	.	.	.	7	.	.	.	.	4	.	.	.	.	15	2	1	
	321	.	.	.	3	.	.	.	.	6	.	.	3	.	1	.	.	2	.	.	15	4	1	
	312	.	.	.	3	.	.	.	3	.	.	6	.	.	.	2	.	.	1	.	15	4	1	
	231	.	1	.	.	.	.	6	.	.	.	2	3	.	.	.	3	.	.	.	15	4	1	
	222	.	1	.	.	.	4	.	.	.	5	.	.	.	4	.	.	.	1	.	15	6	1	32
	3111	.	.	.	3	.	.	.	3	.	3	.	3	.	1	.	1	.	1	.	15	8	1	
	1311	.	.	.	3	.	.	.	4	.	1	.	3	.	3	.	.	.	1	.	15	8	1	
	2211	.	1	.	.	.	4	.	.	2	1	2	.	2	.	2	.	.	1	.	15	12	1	
	2121	.	1	.	.	2	.	.	4	.	.	4	.	.	2	.	.	2	.	.	15	12	1	
	2112	.	1	.	.	2	.	2	.	.	5	.	.	2	.	2	.	.	1	.	15	12	1	
	1221	.	.	2	.	.	1	2	.	.	5	.	.	2	1	.	.	2	.	.	15	12	1	
	21111	.	1	.	.	2	.	2	.	2	1	2	1	.	2	.	1	.	1	.	15	24	1	
	12111	.	.	2	.	.	2	.	1	2	1	2	.	2	1	.	1	.	1	.	15	24	1	
	11211	.	1	.	.	2	.	2	1	.	3	.	1	2	.	2	.	.	1	.	15	24	1	
	111111	.	1	.	1	.	2	.	2	.	3	.	2	.	2	.	1	.	1	.	15	48	1	
33	33	1	.	.	.	.	.	9	.	.	.	.	.	9	.	.	.	.	.	1	20	1	1	
	411	.	.	.	2	.	.	.	.	3	.	3	.	.	.	.	2	.	.	.	10	3	2	31
	141	.	.	.	.	3	.	.	.	.	4	.	.	.	.	3	.	.	.	.	10	3	2	
	321	1	.	.	.	.	6	.	.	3	.	3	.	.	6	.	.	.	.	1	20	3	1	
	312	1	.	.	.	3	.	.	6	.	.	.	6	.	.	3	.	.	.	1	20	3	1	
	231	.	.	3	.	.	.	1	6	.	.	.	6	1	.	.	.	3	.	.	20	3	1	
	222	.	1	.	.	.	2	.	.	.	4	.	.	.	2	.	.	.	1	.	10	9	2	
	3111	1	.	.	.	3	.	3	.	3	.	3	.	3	.	3	.	.	.	1	20	6	1	
	1311	.	.	3	.	.	.	4	.	3	.	3	.	4	.	.	.	3	.	.	20	6	1	
	2211	.	2	.	.	1	2	1	.	4	.	4	.	1	2	1	.	.	2	.	20	9	1	
	2121	1	.	.	2	.	.	5	.	.	4	.	.	5	.	.	2	.	.	1	20	9	1	
	2112	1	.	1	.	.	4	.	.	4	.	4	.	.	4	.	.	1	.	1	20	9	1	
	1221	1	.	.	.	4	.	.	4	1	.	1	4	.	.	4	.	.	.	1	20	9	1	
	21111	1	.	1	.	1	2	1	2	.	4	.	2	1	2	1	.	1	.	1	20	18	1	2
	12111	1	.	.	2	.	2	1	2	2	.	2	2	1	2	.	2	.	.	1	20	18	1	
	11211	.	1	.	.	1	.	1	1	.	2	.	1	1	.	1	.	.	1	.	10	36	2	
	111111	1	.	1	.	2	.	3	.	3	.	3	.	3	.	2	.	1	.	1	20	36	1	
411	411	.	.	.	.	.	12	.	.	.	4	.	8	.	4	.	.	1	.	1	30	1	1	
	141	.	.	.	4	.	4	.	.	1	12	1	.	.	4	.	4	.	.	.	30	1	1	
	321	.	.	.	6	.	.	.	6	.	6	3	.	3	2	.	2	.	2	.	30	2	1	
	312	.	.	.	6	.	.	3	.	3	6	.	6	.	2	.	2	.	2	.	30	2	1	
	231	.	2	.	.	.	6	.	6	.	2	.	8	.	.	3	.	3	.	.	30	2	1	
	222	.	1	.	.	2	.	2	.	2	1	2	.	2	.	2	.	.	1	.	15	6	2	

Table 4 (*cont.*)

See explanation on page 166

$\frac{P_1}{P}$	$\frac{P_2}{P}$	{u}	{v}	P(S)																					ΣP	$\frac{P_1}{P}$	$\frac{P_2}{P}$
				− 11	− 10	− 9	− 8	− 7	− 6	− 5	− 4	− 3	− 2	− 1	0	+ 1	+ 2	+ 3	+ 4	+ 5	+ 6	+ 7	+ 8	+ 9	+ 10	+ 11			
1	1	411	3111	.	.	.	.	.	6	.	.	3	.	6	.	6	.	4	.	2	.	2	.	1	.	.	30	4	1
5	2		1311	.	.	.	.	3	.	3	.	1	6	2	.	4	.	6	.	3	.	1	.	1	.	.	30	4	1
10	3		2211	.	.	.	2	.	.	4	.	4	2	.	6	.	4	.	4	.	2	1	.	1	.	.	30	6	1
5	1		2121	.	.	.	2	.	2	.	2	4	.	4	4	.	4	1	2	1	2	.	2	.	.		30	6	1
5	1		2112	.	.	.	2	.	2	.	4	.	2	5	.	5	2	.	4	.	2	.	2	.	.	.	30	6	1
10	1																												
			1221	.	.	.	2	.	2	.	4	.	2	5	.	5	2	.	4	.	2	.	2	.	.	.	30	6	1
10	1		21111	.	.	.	2	.	2	.	4	.	4	1	4	2	2	3	.	3	.	2	.	1	.	.	30	12	1
10	1		12111	.	.	.	2	.	2	1	2	2	2	2	4	1	4	1	2	2	.	2	.	1	.	.	30	12	1
30	2		11211	.	.	1	.	1	2	.	4	1	2	3	2	3	2	1	4	.	2	1	.	1	.	.	30	12	1
20	1		111111	.	.	1	.	2	.	3	.	4	.	5	.	5	.	4	.	3	.	2	.	1	.	.	30	24	1
20	1																												
30	1	141	141	.	.	1	.	.	.	.	8	.	.	.	12	.	.	.	8	.	.	.	.	1	.	.	30	1	1
			321	.	.	.	.	3	.	.	6	.	2	.	8	.	2	.	6	.	.	3	.	.	.	.	30	2	1
30	1		312	.	.	.	.	.	6	.	.	3	2	.	8	.	2	3	.	.	6	.	.	.	.	.	30	2	1
30	1		231	.	.	.	2	.	.	.	6	3	.	.	8	.	.	3	6	.	.	.	2	.	.	.	30	2	1
30	1		222	.	.	.	.	2	.	.	.	4	.	.	3	.	.	4	.	.	.	2	.	.	.	.	15	6	2
60	1		3111	.	.	.	.	3	.	3	.	4	.	2	6	2	.	4	.	3	.	3	.	.	.	.	30	4	1
60	1		1311	.	.	1	.	1	.	3	.	6	.	1	6	1	.	6	.	3	.	1	.	1	.	.	30	4	1
60	1																												
120	1		2211	.	.	.	2	.	2	.	2	4	2	1	4	1	2	4	2	.	2	.	2	.	.	.	30	6	1
			2121	.	.	.	2	.	.	5	.	.	6	.	4	.	6	.	.	5	.	.	2	.	.	.	30	6	1
1	1		2112	.	.	.	.	4	.	.	4	.	4	1	4	1	4	.	4	.	.	4	.	.	.	.	30	6	1
4	3		1221	.	.	1	.	.	4	.	.	4	4	.	4	.	4	4	.	.	4	.	.	1	.	.	30	6	1
2	1		21111	.	.	.	2	.	2	1	2	2	2	3	2	3	2	2	2	1	2	.	2	.	.	.	30	12	1
2	1		12111	.	.	1	.	1	2	1	2	1	4	2	2	2	4	1	2	1	2	1	.	1	.	.	30	12	1
4	1		11211	.	.	1	.	2	.	1	4	.	4	2	2	2	4	.	4	1	.	2	.	1	.	.	30	12	1
4	1																												
		321	321	.	.	3	.	.	.	12	.	.	12	3	.	3	12	.	6	.	6	.	2	.	.	1	60	1	1
4	1		312	.	.	3	.	.	6	.	.	15	.	.	12	.	6	3	6	.	6	1	.	.	2	.	60	1	1
6	1		231	.	.	.	6	.	.	3	8	.	.	13	6	.	6	3	6	.	6	.	.	3	.	.	60	1	1
8	1		222	.	.	2	.	.	2	2	.	2	5	.	4	.	5	2	.	2	2	.	.	2	.	.	30	3	2
8	1		3111	.	.	3	.	.	6	.	6	3	6	6	.	9	.	9	.	6	.	4	.	1	.	1	60	2	1
12	1		1311	.	.	3	.	.	6	3	.	9	.	6	6	4	6	4	.	7	.	3	.	3	.	.	60	2	1
12	1		2211	.	.	4	.	.	6	.	2	9	.	9	.	8	4	4	4	1	4	1	2	.	2	.	60	3	1
12	1		2121	.	2	.	.	5	.	4	4	4	4	5	2	9	2	4	4	4	2	.	4	.	.	1	60	3	1
12	1		2112	.	2	.	2	1	2	5	2	4	4	8	.	8	4	4	2	5	2	1	2	.	2	.	60	3	1
24	1		1221	1	.	.	2	4	.	4	2	9	2	.	12	.	2	9	2	4	.	4	2	.	.	1	60	3	1
24	1		21111	.	2	.	2	1	4	1	6	1	8	1	8	1	8	2	4	3	2	3	.	2	.	1	60	6	1
24	1		12111	1	.	.	4	.	4	.	8	1	6	3	6	4	4	3	6	1	4	1	2	1	.	1	60	6	1
48	1		11211	.	2	.	2	2	2	3	4	2	8	1	8	1	8	2	4	3	2	2	2	.	2	.	60	6	1
			111111	1	.	2	.	4	.	6	.	8	.	9	.	9	.	8	.	6	.	4	.	2	.	1	60	12	1
1	1																												
3	2	312	312	.	.	3	.	3	.	.	12	.	.	12	6	.	6	.	12	.	.	3	2	.	.	1	60	1	1
3	2		231	.	.	3	.	.	6	1	.	12	2	.	12	.	6	3	6	3	.	.	6	.	.	.	60	1	1
3	1		222	.	1	.	.	2	1	2	.	2	5	.	4	.	5	2	.	2	1	2	.	.	1	.	30	3	2
3	1		3111	.	.	3	.	3	.	3	6	3	6	3	6	6	.	9	.	6	.	4	.	1	.	1	60	2	1
3	1		1311	.	.	.	6	.	.	6	.	9	.	9	.	7	6	1	6	4	.	3	.	3	.	.	60	2	1
			2211	.	2	.	.	5	.	5	2	4	4	8	.	9	2	5	4	.	6	.	2	.	2	.	60	3	1
9	2																												
6	1		2121	1	.	.	4	.	.	9	.	4	4	8	2	4	6	4	2	5	2	1	2	.	2	.	60	3	1
6	1		2112	1	.	1	2	.	6	.	2	8	2	8	.	8	2	8	2	.	6	.	2	1	.	1	60	3	1
9	1		1221	.	2	.	.	5	.	4	4	4	4	5	4	5	4	4	4	4	.	5	.	.	2	.	60	3	1
9	1		21111	1	.	1	2	1	4	1	6	.	10	.	8	1	8	2	4	3	2	3	.	2	.	1	60	6	1
9	1		12111	.	2	.	2	1	4	1	6	2	6	3	6	3	6	2	6	1	4	1	2	1	.	1	60	6	1
			11211	.	2	.	2	1	4	2	4	3	6	2	8	2	6	3	4	2	4	1	2	.	2	.	60	6	1
9	1																												
18	1	231	231	.	.	.	6	.	.	.	14	.	.	3	12	3	.	12	.	.	6	3	.	.	.	1	60	1	1
18	1		222	.	1	.	.	.	5	.	.	4	1	2	4	2	1	4	.	.	5	.	.	.	1	.	30	3	2
36	2		3111	.	.	3	.	3	.	4	6	1	6	4	6	6	.	9	.	6	.	3	.	3	.	.	60	2	1
36	1		1311	1	.	.	.	6	.	3	6	.	6	9	.	9	.	4	6	4	.	3	.	3	.	.	60	2	1
			2211	.	2	.	.	4	2	4	4	.	8	4	2	9	2	5	2	4	2	4	.	1	.	1	60	3	1
1	1		2121	.	.	4	.	.	6	.	4	5	4	5	2	8	4	4	.	9	.	.	4	.	.	1	60	3	1
1	1																												
2	1		2112	.	2	.	.	5	.	5	2	4	6	4	4	4	6	4	2	5	.	5	.	.	2	.	60	3	1
2	1		1221	.	2	.	.	5	2	.	6	4	2	9	.	9	2	4	6	.	2	5	.	.	2	.	60	3	1
2	1		21111	.	2	.	2	1	4	2	4	3	6	3	6	2	8	1	6	1	4	1	2	1	.	1	60	6	1
6	2		12111	1	.	1	2	1	4	1	6	1	8	1	8	2	6	3	4	3	2	2	2	.	2	.	60	6	1
			11211	1	.	1	2	.	6	1	4	3	6	2	8	2	6	3	4	1	6	.	2	1	.	1	60	6	1

Table 4 (*cont.*)

See explanation on page 166

{u}	{v}	P(S)																													ΣP	$\frac{P_1}{P}$	$\frac{P_2}{P}$
		−15	−14	−13	−12	−11	−10	−9	−8	−7	−6	−5	−4	−3	−2	−1	0	+1	+2	+3	+4	+5	+6	+7	+8	+9	+10	+11	+12	+13	+14	+15			
222	222	.	.	.	1	.	.	.	8	.	.	8	2	16	.	.	20	.	.	16	2	8	.	.	8	.	.	.	1	.	.	.	90	1	1
	3111	.	.	.	.	.	1	.	1	.	3	.	2	2	2	4	.	4	2	2	2	.	3	.	1	.	1	.	.	.	.	.	30	4	3
	1311	.	.	.	.	.	1	.	.	2	1	2	2	.	4	2	2	2	4	.	2	2	1	2	.	.	1	.	.	.	.	.	30	4	3
	2211	.	.	.	1	.	.	2	4	2	4	.	14	.	8	8	4	8	8	.	14	.	4	2	4	2	.	.	1	.	.	.	90	2	1
	2121	.	.	.	.	1	.	.	3	.	2	3	2	2	6	.	7	.	6	2	2	3	2	.	3	.	.	1	.	.	.	.	45	4	2
	2112	.	.	.	1	.	.	4	.	4	4	.	14	.	8	4	12	4	8	.	14	.	4	4	.	4	.	.	1	.	.	.	90	2	1
	1221	.	.	.	.	.	4	.	.	8	.	.	15	.	8	4	12	4	8	.	15	.	.	8	.	.	4	.	.	.	.	.	90	2	1
	21111	.	.	.	1	.	1	2	2	4	1	8	2	10	1	12	2	12	1	10	2	8	1	4	2	2	1	.	1	.	.	.	90	4	1
	12111	.	.	.	.	2	.	2	2	4	2	6	3	10	1	12	2	12	1	10	3	6	2	4	2	2	.	2	.	.	.	.	90	4	1
	11211	.	.	.	1	.	.	4	.	6	.	8	2	10	3	8	6	8	3	10	2	8	.	6	.	4	.	.	1	.	.	.	90	4	1
	111111	.	.	.	1	.	2	.	5	.	7	.	11	.	12	.	14	.	12	.	11	.	7	.	5	.	2	.	1	.	.	.	90	8	1
3111	3111	.	.	.	.	.	.	6	.	6	.	12	.	18	3	12	9	6	15	.	15	.	10	.	5	.	2	.	1	.	.	.	120	1	1
	1311	.	.	.	.	.	.	6	.	6	3	6	9	6	12	6	10	12	5	12	5	6	7	.	6	.	3	.	.	.	.	.	120	1	1
	2211	.	.	.	.	.	2	.	2	1	4	2	4	2	8	1	8	2	6	4	2	5	.	4	.	2	.	1	.	.	.	.	60	3	2
	2121	.	.	.	.	1	.	1	2	1	4	1	6	2	6	3	6	3	6	3	4	3	2	3	.	2	.	1	.	.	.	.	60	3	2
	2112	.	.	.	.	1	.	2	.	3	2	3	4	2	8	1	8	1	8	2	4	3	2	3	.	2	.	1	.	.	.	.	60	3	2
	1221	.	.	.	.	1	.	1	2	1	4	1	6	3	4	5	4	5	4	3	6	1	4	1	2	1	.	1	.	.	.	.	60	3	2
	21111	.	.	.	.	2	.	4	.	8	.	12	.	16	.	18	1	16	3	12	5	6	6	2	5	.	3	.	1	.	.	.	120	3	1
	12111	.	.	.	.	2	.	4	.	8	.	12	1	14	3	14	5	12	5	12	3	10	2	6	2	2	2	.	1	.	.	.	120	3	1
	11211	.	.	.	.	2	.	4	1	6	3	8	4	12	3	16	2	16	3	12	4	8	3	6	1	4	.	2	.	.	.	.	120	3	1
	111111	.	.	.	1	.	3	.	6	.	10	.	14	.	17	.	18	.	17	.	14	.	10	.	6	.	3	.	1	.	.	.	120	6	1
1311	1311	.	.	.	.	.	3	.	6	.	4	12	2	12	7	6	12	12	6	12	3	6	9	.	6	.	1	.	1	.	.	.	120	1	1
	2211	.	.	.	.	1	.	1	2	.	6	1	4	3	6	3	6	4	4	5	2	3	4	1	2	1	.	1	.	.	.	.	60	3	2
	2121	.	.	.	.	1	.	1	2	1	4	1	6	2	6	3	6	4	4	5	2	4	2	2	2	.	2	.	.	.	.	.	60	3	2
	2112	.	.	.	.	.	2	.	2	1	4	2	4	4	4	5	4	5	4	4	4	2	4	1	2	.	2	.	.	.	.	.	60	3	2
	1221	.	.	.	.	1	.	1	2	2	2	3	4	3	6	2	8	2	6	3	4	3	2	2	2	1	.	1	.	.	.	.	120	3	1
	21111	.	.	.	.	2	.	4	.	8	.	12	1	14	3	14	5	12	6	10	5	8	3	6	1	4	.	2	.	.	.	.	120	3	1
	12111	.	.	.	.	2	.	4	1	6	3	8	4	12	3	16	2	16	2	14	2	10	2	6	2	2	2	.	1	.	.	.	120	3	1
	11211	.	.	.	1	.	2	2	1	8	1	10	3	12	3	16	2	16	3	12	3	10	1	8	1	2	2	.	1	.	.	.	120	3	1

Table 4 (*cont.*)

See explanation on p. 166

| | | $P(S)$ | | |

Biometrika (1960), **47**, 1 *and* 2, *p.* 173
Printed in Great Britain

Miscellanea

On normalizing the incomplete beta-function for fitting to dose-response curves

BY M. E. WISE
Radiological Protection Service, Belmont, Sutton, Surrey

SUMMARY

It is shown that if the probability integral of x is the incomplete beta-function $I_x(p,q) = \alpha$, the cube root of $-\log x$ is distributed normally to good approximation for all $q \geqslant 2$ and $p \geqslant q$, at least for probabilities α between 0·05 and 0·95. This makes it easy to fit dose-response curves when α is the probability of response and x depends on the dose, and to do so more simply than from another transformation of x recently published in *Biometrika*.

1. INTRODUCTION

Recently Kimball & Leach (1959) put forward an interesting application of the incomplete beta distribution to a radiation dose-response problem. The instance quoted could arise in 'multiple target theory'. Given a finite number of targets, the distribution of the number of hits for a given dose of radiation can be binomial. If response occurs only after at least q hits, the probability α of response is of the form

$$\alpha = I_x(p,q) = \int_0^x t^{p-1}(1-t)^{q-1}\,dt \Big/ \int_0^1 t^{p-1}(1-t)^{q-1}\,dt,$$

where x is a function of the dose. There are, of course, many other situations that could give rise to a similar relationship.

The authors' problem was to replace x by a normally distributed variate, by means of a transformation of x not involving p and q. Then the normal deviate y_α corresponding to α, such that

$$\frac{1}{\sqrt{2\pi}}\int_{y_\alpha}^{\infty} \exp -\tfrac{1}{2}t^2\,dt = \alpha$$

can be written

$$y_\alpha = A + BX,$$

where X is the transformed function of x and A and B are functions of p and q. As they point out, this is just the form of relationship needed for a probit analysis of the dose-response data.

The transformation given is rather complex, namely

$$X^3 = \frac{1-x}{1+x}\left(\frac{21}{23} + \frac{1}{19x} + \frac{1}{3}\exp -\frac{5}{2x}\right).$$

Their numerical tables show that it achieves its aim except when q is between about $0{\cdot}8p$ and $0{\cdot}9p$, when the error is often large. But the results present a challenge to find something better or simpler.

The improved solution is very simple, namely

$$X^3 = -\log_e x.$$

2. THE BASIS OF THE TRANSFORMATION

The only difference from the method of Kimball & Leach arises in relating $I_x(p,q)$ to the χ^2 distribution; afterwards χ^2 is normalized in the same way as before. If $\chi^2_{2q}(\alpha)$ is the χ^2 value with $2q$ degrees of freedom and a probability α of being exceeded, we can put

$$\tfrac{1}{2}\chi^2_{2q}(\alpha) = -N\log_e x\left[1 - \frac{(q-1)\{q+1+\tfrac{1}{2}\chi^2_{2q}(\alpha)\}}{24N^2} + O(N^{-4})\right]\ldots, \tag{1}$$

where

$$N = p + \tfrac{1}{2}q - \tfrac{1}{2} \quad \text{and} \quad \alpha = I_x(p,q).$$

The N^{-4} term in the bracket is about 1/400 when $p \doteqdot q$ and is proportional to $(q/N)^4$. One should choose $p \geqslant q$, if not, p and q can be interchanged since $1-\alpha = I_{1-x}(q,p)$. For details, see Wise (1950).

In order to 'separate the variables', the χ^2 term in the bracket in (1) must be replaced by an approximation to it independent of α and x. In terms of the normal deviate y_α (Campbell, 1923; Riordan, 1949)

$$\tfrac{1}{2}\chi^2_{2q}(\alpha) = q - \tfrac{1}{3} + y_\alpha q^{\frac{1}{2}} + \tfrac{1}{3}y_\alpha^2 + \tfrac{1}{36}(y_\alpha^3 - 7y_\alpha)\,q^{-\frac{1}{2}} + \ldots. \tag{2}$$

Hence we can put $\frac{1}{2}\chi^2_{2q}(\alpha)$ equal to q or to $q-\frac{1}{3}$. The latter has the advantage of being almost exactly right at the 50 % point. This approximation is a poor one for small q and extreme probabilities (very near 0 or 1), but the error is in all cases multiplied by the small factor $(q-1)/(24N^2)$. A related approximation was in fact chosen by Hartley & Fitch (1951) as the basis of their charts, and these give the incomplete beta-function 'to graphical accuracy'.

For normalizing χ^2 the Wilson–Hilferty (1931) transformation is again used

$$\left\{\frac{\chi^2_{2q}(\alpha)}{2q}\right\}^{\frac{1}{3}} \doteqdot 1 - \frac{1}{9q} + \frac{y_\alpha}{3\sqrt{q}}. \tag{3}$$

This distribution of the cube root of χ^2 has coefficients of skewness $\gamma_1 = 4/(27q^{\frac{1}{2}})$ and of excess $\gamma_2 = -2/(9q)$, respectively. (See, for example, Kendall & Stuart, 1958, § 16·7.)

Substituting in (1) the approximation from (3) and the one obtained from putting $y_\alpha = 0$ in (2) yields

$$y_\alpha \doteqdot -3\sqrt{q}\left(1-\frac{1}{9q}\right) + 3(-\log_e x)^{\frac{1}{3}}\left[N\sqrt{q}\left\{1-\frac{(q-1)\,(q+\frac{1}{3})}{12N^2}\right\}\right]^{\frac{1}{3}}. \tag{4}$$

Hence $(-\log_e x)^{\frac{1}{3}}$ has mean $\quad q^{\frac{1}{3}}\left(1-\dfrac{1}{9q}\right)\left\{N-\dfrac{(q-1)\,(q+\frac{1}{3})}{12N}\right\}^{-\frac{1}{3}}$

and standard deviation nearly proportional to $q^{-\frac{1}{6}}N^{-\frac{1}{3}}$.

3. Numerical accuracy

The largest source of error in (4) is generally the error in the normalizing approximation (3); for any one q and p this varies in quite a complex way over the range of probabilities α from 0 to 1. For a fixed q/p the error slowly decreases as q increases; when q only is fixed, it decreases with q/p because the χ^2 term on the right of (1) has been replaced by $q-\frac{1}{3}$. At any one probability level α, of course, the two sources of error act sometimes in the same and sometimes in opposite directions. Table 1 shows exact and calculated values of α for some small values of p and q. (The x values substituted in (4) were the exact percentage points as tabulated by Thompson (1941) corresponding to $\alpha = 0{\cdot}05$, $0{\cdot}25$, $0{\cdot}50$, $0{\cdot}75$ and $0{\cdot}95$.) The values of p and q were the smaller ones chosen for tabulation by Kimball & Leach, except for $p = 2$, $q = 1$. The errors are of the same order of magnitude as for their approximations, but vary rather more regularly.

Table 1. *Values of $I_x(p,q) = \alpha$, calculated from equation (4); x values obtained from Thompson's tables are such that the exact α are 0·05, 0·25, etc.*

$q \rightarrow$	2	2	2	4	6	10
$p \rightarrow$	2	5	12	5	12	12
Exact α			Calculated values of α			
0·05	0·0456	0·0486	0·0492	0·0464	0·0483	0·0464
0·25	0·2453	0·2482	0·2489	0·2460	0·2479	0·2454
0·50	0·5021	0·5020	0·4979	0·5003	0·5003	0·4993
0·75	0·7555	0·7532	0·7534	0·7531	0·7520	0·7528
0·95	0·9483	0·9477	0·9486	0·9503	0·9503	0·9514

The comparison with their approximation $\hat{I}$ is of most interest, as this is the one that can be used in the dose-response problem. Table 2 shows this comparison for the same set of p and q, and with the values of x used by them. The errors in (4) are more often the smaller, although not always.

For extreme values of α, nearer to 0 or 1, both approximations will generally have larger relative errors than the ones tabulated, but these extremes are less important in dose-response analysis. At any rate it seems clear that over at least 90 % of the incomplete β distribution the cube-root log transformation effectively normalizes it.

We have also given in Table 2 the errors for Kimball & Leach's approximation I^*, although as they have pointed out this approximation cannot be used for 'linearization'.

Table 2. *Comparison of* (4) *with Kimball & Leach's approximations* $\hat{I}$ *and* I^*.
I = *exact value of* $I_x(p, q)$; I_N = *approximation* (4)†

$q \rightarrow$	2	2	4	2	6	10
$p \rightarrow$	2	5	5	12	12	12
x	0·14	0·42	0·29	0·68	0·48	0·37
$10^4\, I$	0533	0510	0505	0473	0517	0481
$10^4\,(\hat{I}-I)$	− 85	+15	−34	+15	+03	− 53
$10^4\,(I_N-I)$	− 47	−15	−36	−07	−18	− 35
$10^4\,(I^*-I)$	− 21	−10	−16	−07	−07	− 11
x	0·36	0·64	0·47	0·82	0·61	0·49
$10^4\, I$	2955	3006	2999	2920	2923	2989
$10^4\,(\hat{I}-I)$	− 74	+31	−81	+40	+04	−152
$10^4\,(I_N-I)$	− 36	−10	−35	−06	−18	− 41
$10^4\,(I^*-I)$	+121	+20	+36	−02	+08	+ 17
x	0·64	0·82	0·65	0·92	0·73	0·60
$10^4\, I$	7045	7044	7064	7206	7011	6914
$10^4\,(\hat{I}-I)$	− 22	+50	−42	+59	+15	−119
$10^4\,(I_N-I)$	+ 46	+35	+30	+32	+19	+ 22
$10^4\,(I^*-I)$	+ 59	+39	+30	+32	+16	+ 16
x	0·86	0·94	0·81	0·97	0·83	0·71
$10^4\, I$	9467	9541	9524	9436	9452	9452
$10^4\,(\hat{I}-I)$	− 35	−23	−19	−16	−02	− 33
$10^4\,(I_N-I)$	− 15	−25	+03	−21	+04	+ 15
$10^4\,(I^*-I)$	− 30	+27	−13	−22	−05	− 07

† I would like to thank Dr Kimball for sending full details of his results.

REFERENCES

CAMPBELL, G. A. (1923). Probability curves showing Poisson's exponential summation. *Bell Syst. Tech. J.* **2**, 97.

HARTLEY, H. O. & FITCH, E. R. (1951). A chart for the incomplete beta function and the cumulative binomial distribution. *Biometrika*, **39**, 423.

KENDALL, M. G. & STUART, A. (1958). *Advanced Theory of Statistics*, **1**. London: Chas. Griffin and Co. Ltd.

KIMBALL, A. W. & LEACH, E. (1959). Approximate linearization of the incomplete β function. *Biometrika*, **46**, 214.

RIORDAN, JOHN (1949). Inversion formulas in normal variable mapping. *Ann. Math. Statist.* **20**, 417.

THOMPSON, CATHERINE M. (1941). Tables of percentage points of the incomplete beta-function. *Biometrika*, **32**, 151 and in *Biometrika Tables for Statisticians*, **1**. Cambridge: the University Press for the Biometrika Trust.

WILSON, E. B. & HILFERTY, M. M. (1931). The distribution of chi-square. *Proc. Nat. Acad. Sci., Wash.*, **17**, 684.

WISE, M. E. (1950). The incomplete beta function as a contour integral and a quickly converging series for its inverse. *Biometrika*, **37**, 208.

On certain functions of normal variates which are uncorrelated of a higher order

BY R. G. LAHA AND E. LUKACS*

The Catholic University of America, Washington, D.C.

It is well known that two linear forms $L_1 = a_1x_1 + \ldots + a_nx_n$ and $L_2 = b_1x_1 + \ldots + b_nx_n$ in independently and normally distributed variates $x_1, \ldots, x_n$ are uncorrelated if, and only if, the relation

$$\mathbf{ab}' = 0 \qquad (1a)$$

* This work was supported by the National Science Foundation.

is satisfied. Moreover (1*a*) is equivalent to the stochastic independence of L_1 and L_2. Thus the condition for the stochastic independence of two linear forms can be expressed either in terms of the coefficients as equation (1*a*) or in terms of the product moment of the forms as

$$E(L_1 L_2) = E(L_1)\,E(L_2). \tag{1b}$$

A number of authors investigated the independence of two homogeneous forms $Q_1 = \mathbf{xAx'}$ and $Q_2 = \mathbf{xBx'}$ in independently distributed normal variates with zero mean and unit variances. We mention here only Craig (1943), Hotelling (1944) and Matusita (1949). These authors showed that Q_1 and Q_2 are independent if, and only if,

$$\mathbf{AB} = 0. \tag{2a}$$

A different condition, in terms of product moments, was given by Kawada (1950) who proved that the four relations

$$E(Q_1^i Q_2^j) = E(Q_1^i)\,E(Q_2^j) \quad (i = 1, 2;\ j = 1, 2) \tag{2b}$$

are equivalent to the condition (2*a*). Kawada's condition can be formulated more concisely by introducing the following terminology.

Let x and y be two random variables and suppose that the expectation $E(x^r y^s)$ exists. We say that x and y are uncorrelated of order (r, s) if the relations

$$E(x^i y^j) = E(x^i)\,E(y^j) \quad (i = 1, \ldots, r;\ j = 1, \ldots, s)$$

hold. In our terminology (2*b*) means Q_1 and Q_2 are uncorrelated of order (2, 2).

In this note we prove the following theorem

THEOREM. Let $x_1, x_2, \ldots, x_n$ be n independent normal variates with zero mean and unit variance. If the inhomogeneous quadratic form $Q = \mathbf{xAx'} + \mathbf{bx'}$ and the linear form $L = \mathbf{cx'}$ are uncorrelated of order (2, 2) then $\mathbf{cA} = 0$ and $\mathbf{cb'} = 0$.

Proof. We note first that $E(x_i^2) = 1$, $E(x_i^4) = 3$, $E(x_i^6) = 15$. Using these values we obtain after some straightforward but tedious computations the following expressions

$$\left.\begin{aligned} E(LQ) - E(L)\,E(Q) &= \mathbf{cb'}, \\ E(L^2Q) - E(L^2)\,E(Q) &= 2\mathbf{cAc'}, \\ E(LQ^2) - E(L)\,E(Q^2) &= 4\mathbf{cAb'} + 2\mathbf{cb'}\,\mathrm{tr}\,\mathbf{A}, \\ E(L^2Q^2) - E(L^2)\,E(Q^2) &= 8\mathbf{cAA'c'} + 4\mathbf{cAc'}\,\mathrm{tr}\,\mathbf{A} + 2(\mathbf{cb'})^2. \end{aligned}\right\} \tag{3}$$

We write here $\mathrm{tr}\,\mathbf{A} = \sum_{i=1}^{n} a_{ii}$ for the trace of the matrix $\mathbf{A}$. Since Q and L are uncorrelated of order (2, 2) conditions (3) reduce to

$$\mathbf{cA} = 0, \quad \mathbf{cb'} = 0 \tag{4}$$

so that the theorem is established.

It was shown by Laha (1956) that conditions (4) ensure the independence of L and Q. Therefore the independence of L and Q follows from the assumption that L and Q are uncorrelated of order (2, 2). The converse statement is obvious and we obtain the following corollary to our theorem.

COROLLARY. Let $x_1, x_2, \ldots, x_n$ be n independent normal variables with zero mean and unit variance. The inhomogeneous quadratic polynomial $Q = \mathbf{xAx'} + \mathbf{bx'}$ and the linear form $L = \mathbf{cx'}$ are independent if, and only if, they are uncorrelated of order (2, 2).

If the polynomial Q is homogeneous, that is if $\mathbf{b} = 0$, then we can conclude from Kawada's result that L and Q are independent if they are uncorrelated of order (4, 2). Our result shows that a weaker assumption is sufficient for the independence of L and Q.

It would be desirable to obtain further results concerning the connexion between the order of uncorrelatedness and the degree of two independent polynomial statistics in independent normal variates.

REFERENCES

CRAIG, A. T. (1943). Note on the independence of certain quadratic forms. *Ann. Math. Statist.* **14**, 195–7.

HOTELLING, H. (1944). Note on a matrix theorem of A. T. Craig. *Ann. Math. Statist.* **15**, 427–9.

KAWADA, Y. (1950). Independence of quadratic forms in normally correlated variables. *Ann. Math. Statist.* **21**, 614–5.

LAHA, R. G. (1956). On the stochastic independence of two second-degree polynomial statistics in normally distributed variables. *Ann. Math. Statist.* **27**, 790–6.

MATUSITA, K. (1949). Note on the independence of certain statistics. *Ann. Inst. Statist. Math., Tokyo*, **1**, 79–82.

A further note on a simple method for fitting an exponential curve

By H. D. PATTERSON

Rothamsted Experimental Station, Harpenden, Herts.

Introduction

The least-squares method for fitting the exponential regression curve

$$\mathscr{E}(y_x) = \alpha - \beta\rho^x, \quad \text{where} \quad 0 < \rho < 1 \quad (x = 0, 1, 2, \ldots, n-1) \tag{1}$$

requires the solution of a polynomial equation of degree $3n-8$ in $\hat{\rho}$, the estimate of ρ. Stevens (1951) described a method for obtaining $\hat{\rho}$ and the least-squares estimates of α and β by successive approximation. Patterson (1956) pointed out that the polynomial equation can be written in the form

$$r = \frac{\sum_{1}^{n-1} w_x(\hat{\rho})\, y_x}{\sum_{1}^{n-1} w_x(\hat{\rho})\, y_{x-1}} \quad (\Sigma w_x = 0), \tag{2}$$

where the $w_x(\hat{\rho})$ are polynomials of degree $3n-9$ and suggested that when full efficiency is not required a useful approximation to $\hat{\rho}$ can be obtained very simply from

$$r' = \frac{\Sigma w_x y_x}{\Sigma w_x y_{x-1}} \quad (\Sigma w_x = 0) \tag{3}$$

with the w_x given by the $w_x(\hat{\rho})$ for some specified value of $\hat{\rho}$; α and β can then be estimated by the regression y_x on r'^x. Suitable values of the w_x leading to estimates of moderate to high efficiency were given for the cases $n = 4$, 5, 6 and 7.

This approximate method can be extended to larger values of n but, as pointed out by Patterson & Lipton (1959), at least two sets of coefficients w_x are required to cover the entire range of ρ with reasonable efficiency. In practice the use of alternative estimates is likely to be inconvenient.

The purpose of the present note is to present a modification of the method which leads to increased efficiency and closer approximations to $\hat{\rho}$ without adding unduly to the arithmetic.

Modified method for four equally spaced levels of x

The modification consists of replacing the w_x of (3) by functions $u_x + rv_x$ chosen to give as good approximations as possible to the polynomials $w_x(r)$; the resulting estimate, which will be denoted simply by r, is then one of the roots of the quadratic equation

$$r^2\Sigma v_x y_{x-1} - r(\Sigma v_x y_x - \Sigma u_x y_{x-1}) - \Sigma u_x y_x = 0. \tag{4}$$

When $n = 4$ the u_x and v_x can be chosen so that

$$u_x + r_i v_x \propto w_x(r_i)$$

for three different values $r_i = r_1, r_2, r_3$. The estimate r is then equal to $\hat{\rho}$ for each of the three values. When the r_i are taken to be 0, 0·5 and 1 the values of u_x and v_x (with arbitrary factor) are

x	u_x	v_x
1	−10	−17
2	5	− 5
3	5	22

(5)

A measure of the degree of approximation involved is provided by the asymptotic efficiency of r when the y_x are independent and have equal variance σ^2. Under these conditions the variance of r tends to

$$\frac{(1+\rho^2)\,\Sigma(u_x+\rho v_x)^2 - 2\rho\Sigma(u_x+\rho v_x)\,(u_{x-1}+\rho v_{x-1})}{\{\Sigma(u_x+\rho v_x)\,\rho^{x-1}\}^2}\,\frac{\sigma^2}{\beta^2} \tag{6}$$

as σ^2/β^2 tends to zero; the corresponding variance of $\hat{\rho}$ is obtained by substituting $w_x(\rho)$ for $u_x + \rho v_x$ in (6) or, more conveniently, from

$$\frac{F_{rr}\sigma^2}{\beta^2},$$

where F_{rr} is a quantity defined by Stevens (1951). The efficiency of r is very high (over 99·9 %) throughout the entire range of ρ. As 100 % efficiency can, with this estimate, only be obtained if $r = \hat{r}$ the approximation to $\hat{r}$ appears to be a good one.

As a further test of this point various inefficient estimates, including r, have been calculated from the data of seven replicated field experiments on the responses of barley to four equally spaced levels of nitrogen fertilizer and compared with the least-squares estimates $\hat{r}$. The estimates considered, apart from r and $\hat{r}$, were

(1) the ratio $$r_{P,1\cdot25} = \frac{4y_3 + y_2 - 5y_1}{4y_2 + y_1 - 5y_0}$$

proposed by Patterson (1956).

(2) r_T, the regression of y_x on y_{x-1}, where $x = 1, 2, 3$.

(3) r_H, the estimate given by the method of Hartley (1948).

(4) r_1, the estimate obtained after one cycle of the iterative method of Stevens (1951), using $r_{P,1\cdot25}$ as the preliminary estimate.

The estimates r_1 and $\hat{r}$ were calculated on a high-speed computer. Estimates (1), (2) and (3) were discussed in considerable detail for the case $n = 4$ by Finney (1958). These three estimates, and especially r_T and r_H, have high asymptotic efficiencies when $n = 4$.

The experimental data and the least-squares estimates $\hat{r}$ are set out in Table 1. Table 2 gives the differences between each of the inefficient estimates and $\hat{r}$. In general r provides a much closer approximation to $\hat{r}$ than $r_{P,1\cdot25}$, r_T or r_H. It is also usually better than r_1 so that replacing $r_{P,1\cdot25}$ by r effectively saves at least one cycle of the Stevens method.

Table 1. *Experimental data and least-squares estimate of* ρ

	Level of nitrogen				
Exp.	$x=0$	1	2	3	$\hat{r}$
	Barley grain yields (cwt./acre)				
1	32·2	38·6	38·6	39·6	0·09476
2	13·5	17·1	18·7	19·1	·39259
3	29·7	34·3	36·7	37·2	·42670
4	15·8	23·4	28·5	32·2	·69191
5	29·0	35·0	37·6	37·7	·33292
6	28·9	36·1	37·5	37·2	·14816
7	26·6	32·1	34·2	34·6	·33750

Table 2. *Differences (multiplied by* 10^5*) between inefficient estimates and* $\hat{r}$ *for data of Table* 1

Exp.	$r-\hat{r}$	$r_{P,1\cdot25}-\hat{r}$	$r_T-\hat{r}$	$r_H-\hat{r}$	$r_1-\hat{r}$
1	69	3024	−1664	2169	533
2	2	85	215	−85	−2
3	2	275	191	−101	−19
4	−2	−191	−76	−1	−4
5	1	−124	190	−226	10
6	−18	−874	150	−465	51
7	0	−45	215	−116	2

Closer approximations than r are rarely likely to be required in practice, at any rate with data of the type considered here. The largest difference between r and $\hat{r}$ occurs with the data for which the fitness of the exponential curve is poorest (Exp. 1). The fact that these data have the smallest value of $\hat{r}$ is coincidental.

Estimation of ρ with more than four equally spaced levels

Equation (4) can also be used to provide estimates of ρ for $n > 4$, but as the degree of the polynomials $w_x(\hat{r})$ increases with n the results cannot be expected to be as good as in the case $n = 4$. Nevertheless, by a suitable choice of u_x and v_x a high standard of efficiency and close approximation to $\hat{r}$ can be maintained, at least up to $n = 12$.

When $n > 4$ the $u_x + r_i v_x$ can be made exactly proportional to the $w_x(r_i)$ for only two values of r_i. This will determine quantities $c_1 u_x$ and $c_2 v_x$; the ratio c_1/c_2 can then be chosen so as to minimize the variance of the estimate for a third value of r_i.

Suitable values for $n = 5$ are as follows

x	u_x	v_x
1	−4·11	−3·45
2	1·27	−5·54
3	1·74	2·50
4	1·10	6·49

(7)

With these values of u_x and v_x the asymptotic efficiency of r is more than 99·9 % between about $\rho = 0{\cdot}12$ and $\rho = 0{\cdot}75$ but falls to 99·0 % for $\rho = 0$ and 99·4 % for $\rho = 1$.

Table 3. *Values of u_x and v_x for $n = 6$ to 12*

$n=6$										
u_x	−6·59	1·03	2·58	1·84	1·14					
v_x	−6·29	−11·71	−1·96	8·87	11·09					
$n=7$										
u_x	−7·88	0·20	2·89	2·44	1·45	0·90				
v_x	−1·96	−10·94	−6·85	2·21	8·85	8·69				
$n=8$										
u_x	−8·81	−0·88	2·85	3·05	1·99	1·09	0·71			
v_x	1·17	−9·02	−8·92	−2·70	4·31	8·21	6·95			
$n=9$										
u_x	−9·47	−2·07	2·49	3·47	2·68	1·54	0·80	0·56		
v_x	3·45	−6·69	−9·33	−5·77	0·08	5·18	7·41	5·67		
$n=10$										
u_x	−9·91	−3·24	1·86	3·62	3·30	2·19	1·14	0·58	0·46	
v_x	4·53	−5·09	−9·33	−7·71	−3·00	2·21	6·04	7·22	5·13	
$n=11$										
u_x	−10·27	−4·35	1·08	3·53	3·76	2·89	1·73	0·82	0·43	0·38
v_x	6·39	−2·24	−7·70	−8·13	−5·26	−1·08	2·86	5·42	5·84	3·90
$n=12$										
u_x	−9·81	−5·40	0·09	3·04	3·90	3·44	2·35	1·30	0·57	0·28	0·24
v_x	6·57	−0·21	−6·19	−7·88	−6·47	−3·29	0·38	3·40	5·16	5·16	3·37

Table 4. *Efficiencies of (i) estimates of ρ obtained from equation (4) using the u_x and v_x of Table 3 and (ii) by Hartley's method*

	$n=6$		$n=12$	
$-(n-1)\ln\rho$	(i)	(ii)	(i)	(ii)
0	99·5	100·0	97·9	100·0
2	100·0	100·0	99·9	100·0
4	99·6	99·0	99·0	99·0
6	99·4	97	99·2	96
8	99·3	93	99·8	91
10	99·2	90	99·2	84
12	99·0	87	97	79
25	97	81	77	59
∞ (i.e. $\rho=0$)	97	81	69	51

For $n = 6$ to 12 the values of u_x and v_x were chosen so as to give almost full efficiency in the region of $\rho^{n-1} = 0{\cdot}216$ and in the region of $\rho^{n-1} = 10^{-5}$. These values are set out in Table 3. The asymptotic efficiencies of the estimates are shown in Table 4 for the two cases $n = 6$ and $n = 12$; in the intermediate cases the efficiencies are of comparable order. For purposes of comparison the corresponding efficiencies

of Hartley's estimate are also given. A feature of the Hartley estimate is its high efficiency relative to other inefficient estimates which have been proposed (Patterson & Lipton, 1959). The present series of estimates compare well with the Hartley estimate in terms of asymptotic efficiency.

If desired the values in Table 3 can be rounded off to one place of decimals provided that this is done so that Σu_x and Σv_x remain zero. The rounding off usually results in some slight additional loss of efficiency.

Estimation of ρ with four levels $x = 0, 1, 2, 4$

Most experiments on the responses of crops to fertilizers are designed with equally spaced levels of the fertilizers. Occasionally, however, the levels are taken to be proportional to $x = 0, 1, 2, 4$. In this case also a useful estimate of ρ can be obtained by solving a quadratic equation. A suitable equation is

$$4(y_2 - y_1)r^2 + (4y_3 + y_1 - 5y_0)r - (4y_4 + y_2 - 5y_1) = 0.$$

When $r = 0{\cdot}3$ or $0{\cdot}7$ the estimate is almost identical with $\hat{\rho}$. The asymptotic efficiencies of r are as follows

ρ	0	0·1	0·2	0·3	0·4	0·5
Efficiency (%)	89·3	96·0	99·4	100·0	99·6	99·4
ρ	0·6	0·7	0·8	0·9	1·0	
Efficiency (%)	99·7	100·0	99·9	99·1	97·8	

Summary

A method is presented for estimating ρ in the equation $y = \alpha - \beta\rho^x$ for up to $n = 12$ equally spaced values of x from the root of a quadratic equation in r with coefficients given by linear functions of the y_x. Unless ρ is very small and n large these estimates have very high efficiencies and are good approximations to the least-squares estimates $\hat{\rho}$. The agreement with $\hat{\rho}$ is particularly close when $n = 4$. Details are also given of a similar method for the case in which $x = 0$, 1, 2 and 4.

I wish to thank Mr F. V. Widdowson for providing the experimental data of Table 1.

References

Finney, D. J. (1958). The efficiencies of alternative estimators for an asymptotic regression equation. *Biometrika*, **45**, 370–88.

Hartley, H. O. (1948). The estimation of non-linear parameters by 'internal least squares'. *Biometrika*, **35**, 32–45.

Patterson, H. D. (1956). A simple method for fitting an asymptotic regression curve. *Biometrics*, **12**, 323–9.

Patterson, H. D. & Lipton, S. (1959). An investigation of Hartley's method for fitting an exponential curve. *Biometrika*, **46**, 281–92.

Stevens, W. L. (1951). Asymptotic regression. *Biometrics*, **7**, 247–67.

Estimation of a parameter in the classical occupancy problem

By COLIN R. BLYTH† and GREGORY L. CURME

University of Illinois, Urbana

1. *Introduction and summary.* The following mathematical model has been used [see, for example, Rashevsky (1955), Sprott (1957), Thomas (1956), Curme & Harkness (1957)], for certain biological and combat situations. Each of the r members of an attacking force independently of the others attacks the ith member of a population of n individuals with probability $1/n$ ($i = 1, 2, \ldots, n$). Each attacker has the same probability p of killing the individual he attacks. Let X be the (random) number of population members surviving the attack. Writing A_i for the event that the ith member of the population survives, we have

$$P(A_1, \ldots, A_k) = \left(1 - \frac{k}{n}p\right)^r \quad (k = 1, 2, \ldots, n).$$

† Work supported by the National Science Foundation.

Then as shown by Feller (1950, p. 69)

$$P(X=x)=\binom{n}{x}\sum_{\nu=0}^{n-x}(-1)^{\nu}\binom{n-x}{\nu}\left(1-\frac{x+\nu}{n}p\right)^{r}\quad(x=0,1,\ldots,n). \tag{1}$$

Note that when $r<n$ and $0<p<1$ this probability is zero for $x=0,1,\ldots,n-r-1$ and positive for $x=n-r,n-r+1,\ldots,n$.

This note is concerned with the problem of estimating p from X. For the case $r\leqslant n$ the uniformly minimum variance unbiased estimate of p is obtained and its variance computed. For the case $r>n$ (in which no unbiased estimate of p exists) this same estimate has uniformly minimum variance among estimates having minimum bias in the neighbourhood of $p=0$.

2. *Factorial moments.* The factorial moments of the random variable X with distribution (1) are easily written down

$$\begin{aligned}
EX^{(k)} &= \sum_{x=0}^{n} x^{(k)}\binom{n}{x}\sum_{\nu=0}^{n-x}(-1)^{\nu}\binom{n-x}{\nu}\left(1-\frac{x+\nu}{n}p\right)^{r}\\
&= n^{(k)}\sum_{x=k}^{n}\sum_{\nu=0}^{k}\binom{n-k}{x-k}(-1)^{\nu}\binom{n-x}{\nu}\left(1-\frac{x+\nu}{n}p\right)^{r}\\
&= n^{(k)}\sum_{x=k}^{n}\sum_{i=x}^{n}\binom{n-k}{x-k}(-1)^{i-x}\binom{n-x}{i-x}\left(1-\frac{i}{n}p\right)^{r}\\
&= n^{(k)}\sum_{i=k}^{n}\sum_{x=k}^{i}(-1)^{i}\left(1-\frac{i}{n}p\right)^{r}(-1)^{x}\binom{n-k}{x-k}\binom{n-x}{i-x}\\
&= n^{(k)}\sum_{i=k}^{n}(-1)^{i}\left(1-\frac{i}{n}p\right)^{r}\binom{n-k}{n-i}\sum_{j=0}^{i-k}(-1)^{j}\binom{i-k}{j}\\
&= n^{(k)}\left(1-\frac{k}{n}p\right)^{r}\quad(k=1,2,\ldots,n).
\end{aligned}$$

That is,

$$EX^{(k)}=n^{(k)}\left(1-\frac{k}{n}p\right)^{r}\quad(k=1,2,\ldots). \tag{2}$$

From (2) the moments of X can be written down using the usual formulae for moments in terms of factorial moments. Sprott (1957) finds the first two moments of X using indicator variables and this method is used by Barton & David (1959) to find the factorial moments. For the purposes of this note it is easiest to work with the factorial moments only.

3. *Completeness of the family of distributions.* The family of distributions (1) with $0<p<1$ is complete; that is,

$$Ef(X)\underset{p}{\equiv}0\quad\text{implies}\quad P\{f(X)=0\}\underset{p}{\equiv}1.$$

The proof of this fact is as follows

$$Ef(X)=\sum_{x=0}^{n}f(x)\binom{n}{x}\sum_{\nu=0}^{n-x}(-1)^{\nu}\binom{n-x}{\nu}\left(1-\frac{x+\nu}{n}p\right)^{r}, \tag{3}$$

which is a degree r polynomial in p. Thus $Ef(X)\equiv 0$ implies that every coefficient of this polynomial must be 0. The constant term is

$$\sum_{x=0}^{n}f(x)\binom{n}{x}\sum_{\nu=0}^{n-x}(-1)^{\nu}\binom{n-x}{\nu}=f(n)=0.$$

The coefficient of p is therefore

$$\sum_{x=0}^{n-1}f(x)\binom{n}{x}\sum_{\nu=0}^{n-x}(-1)^{\nu}\binom{n-x}{\nu}\left\{-\frac{r}{n}(x+\nu)\right\}=0,$$

i.e.

$$\sum_{x=0}^{n-1}f(x)\binom{n}{x}\sum_{\nu=0}^{n-x}(-1)^{\nu}\binom{n-x}{\nu}\nu=0,$$

i.e.

$$\sum_{x=0}^{n-1}f(x)\binom{n}{x}(n-x)\sum_{\nu=1}^{n-x}(-1)^{\nu}\binom{n-x-1}{\nu-1}=0,$$

i.e. $nf(n-1)=0$. And in general, if $f(n)=f(n-1)=\ldots=f(n-k+1)=0$, the coefficient of p^{k} is

$$\sum_{x=0}^{n-k}f(x)\binom{n}{x}\sum_{\nu=0}^{n-x}(-1)^{\nu}\binom{n-x}{\nu}\binom{r}{k}\left(\frac{x+\nu}{n}\right)^{k}=0$$

which implies

$$\begin{aligned}\sum_{x=0}^{n-k} f(x)\binom{n}{x}\sum_{\nu=0}^{n-x}(-1)^{\nu}\binom{n-x}{\nu}\{\nu^{(k)}+a_1(x)\,\nu^{(k-1)}+\ldots+a_k(x)\} \\ = \sum_{x=0}^{n-k} f(x)\binom{n}{x}\sum_{\nu=0}^{n-x}(-1)^{\nu}\binom{n-x}{\nu}\nu^{(k)} \\ = \sum_{x=0}^{n-k} f(x)\binom{n}{x}(n-x)^{(k)}\sum_{\nu=k}^{n-x-k}(-1)^{\nu}\binom{n-x-k}{\nu-k} \\ = (-1)^k n^{(k)} f(n-k) = 0.\end{aligned}$$

We can now conclude that $f(n) = f(n-1) = \ldots = f(n-r) = 0$, so that $P\{f(X) = 0\} \equiv 1$ and the family of distributions of X is complete.

4. *Unbiased estimation of p.* Because of completeness, the problem of unbiased estimation of any function of p can be solved by the method of Lehmann & Scheffé (1950). From (3) we observe that

$$Ef(X) = \sum_{k=0}^{n} c_k\left(1-\frac{k}{n}p\right)^r, \tag{4}$$

where $c_0, c_1, \ldots, c_n$ are constants. Therefore no functions of p other than such linear combinations of $\{1-(k/n)p\}^r$, $k = 0, 1, \ldots, n$ possess unbiased estimates. And from (2) we have

$$E\sum_{k=0}^{n} c_k\frac{X^{(k)}}{n^{(k)}} = \sum_{k=0}^{n} c_k\left(1-\frac{k}{n}p\right)^r. \tag{5}$$

The Lehmann & Scheffé (1950) conclusion now is that $\sum_{k=0}^{n} c_k X^{(k)}/n^{(k)}$ is the unique uniformly minimum variance unbiased estimate of $\sum_{k=0}^{n} c_k\left(1-\frac{k}{n}p\right)^r$, where $c_0, c_1, \ldots, c_k$ are arbitrary constants. To find the variance of this estimate we note that

$$X^{(j)}X^{(k)} = \sum_{t=0}^{\min(j,k)} \frac{j^{(t)}k^{(t)}}{t!}X^{(k+j-t)},$$

which gives

$$\begin{aligned}E\left\{\frac{X^{(k)}}{n^{(k)}}-\left(1-\frac{k}{n}p\right)^r\right\}\left\{\frac{X^{(j)}}{n^{(j)}}-\left(1-\frac{j}{n}p\right)^r\right\} \\ = \sum_{t=0}^{\min(j,k)}\frac{j^{(t)}k^{(t)}}{t!}\frac{n^{(k+j-t)}}{n^{(k)}n^{(j)}}\left(1-\frac{k+j-t}{n}p\right)^r-\left(1-\frac{k}{n}p\right)^r\left(1-\frac{j}{n}p\right)^r.\end{aligned}$$

The variance of $\sum_{k=0}^{n} c_k X^{(k)}/n^{(k)}$ is therefore

$$\begin{aligned}E\left\{\sum_{k=1}^{n} c_k\left[\frac{X^{(k)}}{n^{(k)}}-\left(1-\frac{k}{n}p\right)^r\right]\right\}^2 \\ = \sum_{k=1}^{n}\sum_{j=1}^{n} c_j c_k\left\{\sum_{t=0}^{\min(j,k)}\frac{j^{(t)}k^{(t)}}{t!}\frac{n^{(k+j-t)}}{n^{(k)}n^{(j)}}\left(1-\frac{k+j-t}{n}p\right)^r-\left(1-\frac{k}{n}p\right)^r\left(1-\frac{j}{n}p\right)^r\right\}.\end{aligned} \tag{6}$$

In particular, we are interested in estimating p itself. For there to be an unbiased estimate of p we must be able to find constants $c_0, c_1, \ldots, c_n$ such that

$$\sum_{k=0}^{n} c_k\left(1-\frac{k}{n}p\right)^r \equiv p. \tag{7}$$

Equating coefficients of powers of p, (7) becomes

and

$$\left.\begin{aligned} c_0 &= -(c_1+\ldots+c_n); \\ c_1+2c_2+\ldots+nc_n &= -n/r \\ c_1+2^2c_2+\ldots+n^2c_n &= 0 \\ \vdots\quad &\quad\vdots\quad\quad\vdots\quad\quad\vdots \\ c_1+2^rc_2+\ldots+n^rc_n &= 0. \end{aligned}\right\} \tag{8}$$

If $r > n$ the equations (8) are not solvable, because the last n of them have the unique (their determinant of coefficients is of Van der Monde form and clearly non-zero) solution $0, \ldots, 0$ which does not satisfy the second equation. On the other hand, equations (8) have a unique solution if $r = n$ and many solutions if $r < n$.

When $r \leqslant n$, one solution is to take $c_{r+1} = c_{r+2} = \ldots = c_n = 0$; equations (8) then reduce to the following uniquely solvable form.

$$c_0 = -(c_1 + \ldots + c_n); \tag{9}$$

and

$$\left.\begin{array}{l} c_1 + 2c_2 + \ldots + rc_r = -n/r \\ c_1 + 2^2 c_2 + \ldots + r^2 c_r = 0 \\ \quad\vdots \quad\quad \vdots \quad\quad\quad \vdots \quad\quad \vdots \\ c_1 + 2^r c_2 + \ldots + r^r c_r = 0. \end{array}\right\} \tag{10}$$

The determinant of coefficients of equations (10) is of Van der Monde form, so can be readily evaluated

$$\begin{vmatrix} 1 & 2 & \ldots & r \\ 1 & 2^2 & \ldots & r^2 \\ \vdots & \vdots & & \vdots \\ 1 & 2^r & \ldots & r^r \end{vmatrix} = r! \begin{vmatrix} 1 & 1 & \ldots & 1 \\ 1 & 2 & \ldots & r \\ \vdots & \vdots & & \vdots \\ 1 & 2^{r-1} & \ldots & r^{r-1} \end{vmatrix}$$

$$= r!\, 1^{r-1} 2^{r-2} \ldots (r-1) = [1^r 2^{r-1} 3^{r-2} \ldots r]. \tag{11}$$

Using Cramér's rule, the solution for c_k in equations (10) is the ratio of two determinants of which the denominator is (11) and the numerator is (11) with its kth column replaced by the column $-n/r, 0, \ldots, 0$. This numerator is easily reduced to Van der Monde form and its value found to be

$$(-1)^k \frac{n}{rk} \binom{r}{k} [1^r 2^{r-1} 3^{r-2} \ldots r].$$

Equations (10) therefore have the unique solution

$$c_k = (-1)^k \frac{n}{rk} \binom{r}{k} \quad (k = 1, 2, \ldots, r). \tag{12}$$

The unique uniformly minimum variance unbiased estimate of p for the case $r \leqslant n$ is therefore

$$Z = \sum_{k=1}^{r} (-1)^k \frac{n}{rk} \binom{r}{k} \left[\frac{X^{(k)}}{n^{(k)}} - 1\right]. \tag{13}$$

Another solution of equations (8) is obtained by appending additional equations

$$\begin{array}{l} c_1 + 2^{r+1} c_2 + \ldots + n^{r+1} c_n = 0 \\ \quad\vdots \quad\quad\quad \vdots \quad\quad\quad\quad \vdots \quad\quad \vdots \\ c_1 + 2^n c_2 + \ldots + n^n c_n = 0 \end{array}$$

and then solving the whole set in the same way that (9), (10) were solved. This shows (again for the case $r \leqslant n$) that the uniformly minimum variance unbiased estimate for p is

$$Z^* = \sum_{k=1}^{n} (-1)^k \frac{n}{rk} \binom{n}{k} \left[\frac{X^{(k)}}{n^{(k)}} - 1\right]. \tag{14}$$

Because of completeness, Z and Z^* must coincide with probability 1; this can also be checked directly. We may use whichever form we find convenient.

From (6) the variance of Z is

$$\sum_{j=1}^{r} \sum_{k=1}^{r} (-1)^{k+j} \frac{n^2}{r^2 kj} \binom{r}{k} \binom{r}{j} \left\{ \sum_{t=0}^{\min(j,k)} \frac{j^{(t)} k^{(t)}}{t!} \frac{n^{(k+j-t)}}{n^{(k)} n^{(j)}} \left(1 - \frac{k+j-t}{n} p\right)^r - \left(1 - \frac{k}{n} p\right)^r \left(1 - \frac{j}{n} p\right)^r \right\}$$

$$= \frac{1}{r^2} \sum_{j=1}^{r} \sum_{k=1}^{r} \binom{r}{k} \binom{r}{j} \sum_{t=0}^{\min(j,k)} \frac{(j-1)^{(t-1)} (k-1)^{(t-1)} n^{(k+j-t)}}{t!\,(n-1)^{(k-1)} (n-1)^{(j-1)}} \left(1 - \frac{k+j-t}{n} p\right)^r - \left\{p + \sum_{k=1}^{r} (-1)^k \frac{n}{rk} \binom{r}{k}\right\}^2. \tag{15}$$

For fixed r we have the following limits as $n \to \infty$: (i) The variance (15) of Z tends to $p(1-p)/r$; (ii) the value of Z tends to $(n-X)/r$ with probability 1; (iii) the distribution of $n - X$ tends to binomial (r, p). All these limits are what we would expect since the probability is tending to 1 that r different individuals will be attacked.

An unpleasant feature of the estimate Z is that it can exceed 1 for small values of X, being quite large for $n = r$ large and values of X close to 0. This is only a slight drawback owing to the very small probabilities of such values of X even when p is close to 1. We would in practice modify Z by using instead $\min(Z, 1)$. This modified estimate will be slightly biased and will have slightly smaller variance and slightly smaller expected squared error than Z.

5. *Estimation of p with locally minimum bias.* In this section we will consider the case $r > n$, in which there is no unbiased estimate of p. An estimate of p with expectation

$$g(p) = \sum_{k=0}^{n} c_k \left(1 - \frac{k}{n} p\right)^r$$

is said to have minimum bias in the neighbourhood of p_0 if $g(p)$ and as many of its derivatives as possible agree at p_0 with the function p and its derivatives. The unique (and therefore uniformly minimum variance) estimate $\sum_{k=0}^{n} c_k X^{(k)}/n^{(k)}$ with this property is found by solving the equations

$$g(p_0) = p_0, \quad g'(p_0) = 1, \quad g''(p_0) = 0, \quad \ldots, \quad g^{(n)}(p_0) = 0 \quad \text{for} \quad c_0, c_1, \ldots, c_n.$$

These equations are

$$\left.\begin{aligned}
c_0 + \left(1 - \frac{1}{n} p_0\right)^r c_1 + \left(1 - \frac{2}{n} p_0\right)^r c_2 + \ldots + \left(1 - \frac{n}{n} p_0\right)^r c_n &= p_0; \\
\text{and} \qquad \left(1 - \frac{1}{n} p_0\right)^{(r-1)} c_1 + 2\left(1 - \frac{2}{n} p_0\right)^{r-1} c_2 + \ldots + n\left(1 - \frac{n}{n} p_0\right)^{r-1} c_n &= -n/r \\
\left(1 - \frac{1}{n} p_0\right)^{(r-2)} c_1 + 2^2\left(1 - \frac{2}{n} p_0\right)^{r-2} c_2 + \ldots + n^2\left(1 - \frac{n}{n} p_0\right)^{r-2} c_n &= 0 \\
\vdots \qquad\qquad \vdots \qquad\qquad & \vdots \\
\left(1 - \frac{1}{n} p_0\right)^{r-n} c_1 + 2^n\left(1 - \frac{2}{n} p_0\right)^{r-n} c_2 + \ldots + n^n\left(1 - \frac{n}{n} p_0\right)^{r-n} c_n &= 0.
\end{aligned}\right\} \tag{16}$$

The determinant of coefficients of equations (16) is of Van der Monde form, so can be readily evaluated

$$\begin{aligned}
n!\left\{\left(1 - \frac{1}{n} p_0\right) \ldots \left(1 - \frac{n}{n} p_0\right)\right\}^{r-1} &\begin{vmatrix}
1 & 1 & \ldots & 1 \\
\dfrac{1}{1 - \frac{1}{n} p_0} & \dfrac{2}{1 - \frac{2}{n} p_0} & \ldots & \dfrac{n}{1 - \frac{n}{n} p_0} \\
\vdots & \vdots & & \vdots \\
\dfrac{1^{n-1}}{\left(1 - \frac{1}{n} p_0\right)^{n-1}} & \dfrac{2^{n-1}}{\left(1 - \frac{2}{n} p_0\right)^{n-1}} & \ldots & \dfrac{n^{n-1}}{\left(1 - \frac{n}{n} p_0\right)^{n-1}}
\end{vmatrix} \\
&= n!\left\{\left(1 - \frac{1}{n} p_0\right) \ldots \left(1 - \frac{n}{n} p_0\right)\right\}^{r-1} \frac{1^{n-1} \cdot 2^{n-2} \ldots (n-1)}{\left\{\left(1 - \frac{1}{n} p_0\right) \ldots \left(1 - \frac{n}{n} p_0\right)\right\}^{n-1}} \\
&= \left\{\left(1 - \frac{1}{n} p_0\right) \ldots \left(1 - \frac{n}{n} p_0\right)\right\}^{r-n} 1^n \cdot 2^{n-1} \ldots n.
\end{aligned} \tag{17}$$

Using Cramér's rule the solution for c_k in equations (16) is the ratio of two determinants of which the denominator is (17) and the numerator is (17) with its kth column replaced by the column $-n/r, 0, \ldots, 0$. This numerator can be reduced to Van der Monde form and its value found to be

$$(-1)^k \frac{n}{rk} \binom{n}{k} \left\{\left(1 - \frac{1}{n} p_0\right) \ldots \left(1 - \frac{n}{n} p_0\right)\right\}^{r-n} 1^n \cdot 2^{n-1} \ldots n \left(1 - \frac{k}{n} p_0\right)^{n-r}.$$

Equations (16) thus have the unique solution

$$c_k = (-1)^k \frac{n}{rk} \binom{n}{k} \left(1 - \frac{k}{n} p_0\right)^{n-r} \quad (k = 1, 2, \ldots, n). \tag{18}$$

The unique uniformly minimum variance estimate of p with minimum bias in the neighbourhood of p_0 is therefore

$$Z^*(p_0) = p_0 + \sum_{k=1}^{n} (-1)^k \frac{n}{rk} \binom{n}{k} \left\{\frac{X^{(k)}/n^{(k)} - \{1 - (k/n)\, p_0\}^r}{\{1 - (k/n)\, p_0\}^{r-n}}\right\}. \tag{19}$$

Note that in the special case $p_0 = 0$ this reduces to Z^* of § 4. The special case $p_0 = 1$ must be handled separately—only derivatives up to order $n-1$ can be made to agree. The variance of $Z^*(p_0)$ can be written down from (6).

In the case $r > n$ we may then choose a number p_0 near which we think p lies, and then use the estimate $Z^*(p_0)$. Alternative low bias estimates would be to choose a $g(p)$ which in some sense (such as maximum discrepancy) agrees most closely with the function p, and use the uniformly minimum variance unbiased estimate of that $g(p)$.

As to other estimates of p: a minimax estimate for p can be found numerically for small values of n, r. The method we used (discrete *a priori* distributions) works for all n, r but is too laborious for all but the very smallest values. The value of the maximum likelihood estimate of p for a given value of X involves the location of the maximum value on the range $0 \leqslant p \leqslant 1$ of a degree $r-1$ polynomial in p. For the trivial case $r = 1$ and for $r = 2, 3$ this is easily done (the risk function, squared-error loss, is not particularly good for these maximum likelihood estimates). But for $r > 3$ the maximum would have to be located numerically, which is too laborious.

6. *Example.* A flight of n missiles is attacked by r interceptors, each of which is equipped with infrared guidance. The interceptors are guided to the target area by radar and then each seeks a target. We assume that each interceptor chooses a target at random from among the n missiles, has probability p of destroying it and probability $1-p$ of missing it and passing out of the target area. From the number X of missiles surviving this attack, we wish to estimate p, which measures the effectiveness of the interceptors.

If $n = 5$, $r = 4$, formula (13) gives the uniformly minimum variance unbiased estimate for p; when $X = 2$ this estimate has the value 0·98.

If $n = 4$, $r = 5$, no unbiased estimate of p is possible. Supposing values of p close to 1 to be of most interest, we may decide to use formula (19) with $p_0 = 1$ which gives the uniformly minimum variance estimate with minimum bias in the neighbourhood of $p = 1$; when $X = 1$ this estimate has the value 0·80.

The results of § 2 were obtained by one of the authors working jointly with W. Harkness.

References

Barton, D. E. & David, F. N. (1959). A collector's problem. *Trab. Estad.* **10**, 75.

Curme, G. & Harkness, W. (1957). Probability distributions for air battles. *I.A.W.R.* no. 57.

Feller, W. (1950). *An Introduction to Probability Theory and its Applications.* New York: Wiley.

Lehmann, E. L. & Scheffé, H. (1950). Completeness, similar regions and unbiased estimation. I. *Sankhyā*, **10**, 305.

Rashevsky, N. (1955). Note on a combinatorial problem in mathematical biology. *Bull. Math. Biophys.* **17**, 45.

Sprott, D. A. (1957). Probability distributions associated with direct hits on targets. *Bull. Math. Biophys.* **19**, 163.

Thomas, C. J. (1956). The estimation of bombers surviving an air battle. *A.F.O.O.A.* U-1287, T.M. 49.

Geometry and linear discrimination

By C. W. CLUNIES-ROSS and R. H. RIFFENBURGH

Virginia Polytechnic Institute and University of Hawaii

1. Outline of the problem

The problem of discrimination is a problem of classification, viz. on the basis of a set of measurements on an individual that individual is to be classified into one of a set of populations. Although many ramifications are possible this paper will be concerned only with the situation where there are just two possible populations and both of these populations are multivariate normal with respect to the set of measurements characterizing a random member of the population.

The criterion used in defining the best discriminator depends on the situation; however, for a wide class of situations the criterion used is the maximum of the, possibly weighted, conditional probabilities of misclassification. The weights would be either prior probabilities or loss functions, but not both. Another criterion applicable to certain situations is the sum of the, possibly weighted, conditional probabilities of misclassification. The weights here would be the loss functions if relevant loss functions exist and are known. Discrimination based on these two criteria will be referred to as 'minimax' and

'minisum' discrimination, respectively. This paper will be primarily concerned with minimax discrimination.

For both types of discrimination, and indeed for any reasonable type of discrimination, the best discriminator is one of the surfaces of constant likelihood ratio. When the two populations are both multivariate normal these surfaces correspond to a family of similar and similarly situated hyperconics in the spaces of measurements. Except for special cases the identification of the best member of the family (for minimax discrimination) and its evaluation (for minimax or minisum discrimination) are mathematically intractable.

This paper will investigate the best discriminator subject to the restriction that it is linear in the measurements, i.e. is a hyperplane in the measurement space. Much the same problem was the topic of another paper by the authors (Riffenburgh & Clunies-Ross, 1960). That paper considered the optimum constant associated with such a hyperplane, given an arbitrary vector of direction numbers. The current paper considers the determination of such a vector.

Throughout the paper the following notation will be used: $N(\mathbf{M}, \mathbf{\Sigma})$ means either a multivariate normal population with means $\mathbf{M}$ and dispersion matrix $\mathbf{\Sigma}$ or the corresponding density function. The weights will be W_1, W_2, with $W_1 \geqslant W_2$. The two populations are then

$$N(\mathbf{M}_1, \mathbf{\Sigma}_1), \quad N(\mathbf{M}_2, \mathbf{\Sigma}_2). \tag{1·1}$$

If $\mathbf{\Sigma}_1$, $\mathbf{\Sigma}_2$ are non-singular then there is always a non-singular linear transformation of the measurement space to make the two populations

$$N(\mathbf{O}, \mathbf{I}), \quad N(\mathbf{M}, \mathbf{D}^{-1}), \tag{1·2}$$

where $\mathbf{D}$ is diagonal with diagonal elements (d_i) and $0 < d_1 \leqslant d_2 \leqslant \ldots \leqslant d_k$. (There is no real loss of generality in assuming $\mathbf{\Sigma}_1$, $\mathbf{\Sigma}_2$ to be non-singular.)

2. Geometry and some terminology

The family of hyperconics

$$(\mathbf{X}-\mathbf{M})'\mathbf{\Lambda}(\mathbf{X}-\mathbf{M}) = \theta \tag{2·1}$$

represents the constant likelihood contours of the $N(\mathbf{M}, \mathbf{\Lambda}^{-1})$ distribution. The member of (2·1) through a specific point $\mathbf{Y}$ is

$$(\mathbf{X}-\mathbf{M})'\mathbf{\Lambda}(\mathbf{X}-\mathbf{M}) = (\mathbf{Y}-\mathbf{M})'\mathbf{\Lambda}(\mathbf{Y}-\mathbf{M}). \tag{2·2}$$

The tangent (i.e. tangential hyperplane) at $\mathbf{Y}$ to (2·2) is

$$(\mathbf{X}-\mathbf{M})'\mathbf{\Lambda}(\mathbf{Y}-\mathbf{M}) = (\mathbf{Y}-\mathbf{M})'\mathbf{\Lambda}(\mathbf{Y}-\mathbf{M}),$$

i.e.

$$(\mathbf{X}-\mathbf{Y})'\mathbf{\Lambda}(\mathbf{Y}-\mathbf{M}) = 0. \tag{2·3}$$

This tangent will be referred to as the tangent at $\mathbf{Y}$ to the family (2·1). This is permissible, it always exists, and it is unique except for the case where: $\mathbf{Y} = \mathbf{M}$ in which case any hyperplane: $(\mathbf{X}-\mathbf{M})'\mathbf{B} = 0$ may be regarded as a tangent at $\mathbf{M}$ to the family.

For certain points $\mathbf{Y}(t)$ the tangents at $\mathbf{Y}$ to each of two families of the type (2·1) are identical. Using the notation (and the corresponding assumptions) of (1·1) and (1·2) the families are

$$\left.\begin{aligned} \mathbf{X}'\mathbf{X} &= \theta, \\ (\mathbf{X}-\mathbf{M})'\mathbf{D}(\mathbf{X}-\mathbf{M}) &= \theta. \end{aligned}\right\} \tag{2·4}$$

The tangents at $\mathbf{Y}$ to (2·4) will be

$$(\mathbf{X}-\mathbf{Y})'\mathbf{Y} = 0, \tag{2·5}$$

$$(\mathbf{X}-\mathbf{Y})'\mathbf{D}(\mathbf{Y}-\mathbf{M}) = 0. \tag{2·6}$$

These will be identical if and only if there exists some constant t such that

$$t\mathbf{Y} = \mathbf{D}(\mathbf{Y}-\mathbf{M}), \tag{2·7}$$

which determines a locus for $\mathbf{Y}$, viz.

$$\mathbf{Y}(t) = (\mathbf{D}-t\mathbf{I})^{-1}\mathbf{D}\mathbf{M}. \tag{2·8}$$

This locus will play an important role in what follows and will be referred to as the fully common tangent locus although, strictly, it is the locus of the points of contact of fully common tangents. The tangent corresponding to a particular value of t (or a particular point $\mathbf{Y}$ on the fully common tangent locus) will be referred to as the fully common tangent at t (or at $\mathbf{Y}$).

The fully common tangent locus (2·8) consists of a set of branches; each branch superficially resembles a rectangular hyperbola except that it is not coplanar. On examining the positions of the locus with respect to t the following facts are demonstrable

$$\begin{aligned} t &= 0 && \mathbf{Y}(t) = \mathbf{M}, \\ t &= \pm\infty && \mathbf{Y}(t) = \mathbf{O}, \\ t &= d_i && \mathbf{Y}(t) \text{ is infinite.} \end{aligned}$$

The locus is continuous in t except at $t = d_i$ and has as many branches and asymptotes as there are distinct d_i associated with non-zero m_i.

The ranges $(d_1, d_2), \ldots, (d_{k-1}, d_k)$ of t are each associated with one branch of the locus. The split range $[(d_k, \infty)(-\infty, d_1)]$ of t corresponds to the remaining branch of the locus which passes through **O**, **M**. Further results pertaining to the fully common tangent locus are embodied in the following Lemmata.

LEMMA 1. The necessary and sufficient condition for a fully common tangent at t to pass between **O**, **M** is that t be negative.

Proof. The N and S condition may be expressed as: the left side of (2·5) for **M**, i.e. $(\mathbf{M}-\mathbf{Y})'\mathbf{Y}$, and $-\mathbf{Y}'\mathbf{Y}$ be of opposite sign; i.e. that $(\mathbf{M}-\mathbf{Y})'\mathbf{Y}$ be positive. Substituting $\mathbf{Y}(t)$ from (2·8) the condition becomes that: $\mathbf{M}'(\mathbf{I}-(\mathbf{D}-t\mathbf{I})^{-1}\mathbf{D})(\mathbf{D}-t\mathbf{I})^{-1}\mathbf{D}\mathbf{M}$ be positive. The matrices are diagonal, so that the condition may be written as: $-t\mathbf{M}'(\mathbf{D}-t\mathbf{I})^{-1}\mathbf{D}(\mathbf{D}-t\mathbf{I})^{-1}\mathbf{M}$ be positive; i.e. that t be negative. (Note that $\mathbf{Y}(0) = \mathbf{O}$; $\mathbf{Y}(+\infty) = \mathbf{Y}(-\infty) = \mathbf{M}$.)

LEMMA 2. Any hyperplane passing between **O**, **M** intersects the fully common tangent locus at, at least, one negative t.

Proof. A general hyperplane may be expressed as $\mathbf{X}'\mathbf{B}+C = 0$. An N and S condition for it to pass between **O**, **M** is that: C, $\mathbf{M}'\mathbf{B}+C$ be of opposite sign. Substituting (2·8), the intersections of the hyperplane with the locus are given by the roots of

$$\sum_i \frac{b_i m_i d_i}{d_i - t} + C = 0. \tag{2·9}$$

The left-hand side of (2·9) is equal to C when $t = -\infty$, equal to $\sum_i b_i m_i + C = \mathbf{M}'\mathbf{B}+C$ when $t = 0$. Further it is continuous in t for t negative. Therefore (2·9) has at least one negative root.

LEMMA 3. The fully common tangent at t_0 ($t_0 < 0$) does not intersect the fully common tangent locus at any other negative t.

Proof. Intersections for t (as in (2·7)) where $t \neq t_0$ are given by the roots of

$$\sum_i [m_i^2 d_i^2 (t - t_0)/(d_i - t_0)^2 (d_i - t)] = 0. \tag{2·10}$$

All terms in the summation of (2·10) have the same sign for any negative t.

3. TANGENTS AND PROBABILITIES

Any hyperplane $(\mathbf{X}-\mathbf{X}_0)'\mathbf{B} = 0$ divides the measurement space into two regions. If there is an $N(\mathbf{M}, \mathbf{\Lambda}^{-1})$ distribution over the observations, then the probability that an observation will lie in each of the two regions is given by

$$\int_{(\mathbf{X}-\mathbf{X}_0)'\mathbf{B}>0} \cdots \int N(\mathbf{M}, \mathbf{\Lambda}^{-1})\, d\mathbf{X}; \quad \int_{(\mathbf{X}-\mathbf{X}_0)'\mathbf{B}<0} \cdots \int N(\mathbf{M}, \mathbf{\Lambda}^{-1})\, d\mathbf{X}. \tag{3·1}$$

These may be expressed as

$$\int_{-\infty}^{a} N(0,1)\, dx; \quad \int_{a}^{\infty} N(0,1)\, dx, \tag{3·2}$$

where

$$a = (\mathbf{M}-\mathbf{X}_0)'\mathbf{B}/(\mathbf{B}'\mathbf{\Lambda}^{-1}\mathbf{B})^{\frac{1}{2}}. \tag{3·3}$$

LEMMA 4. The hyperplane through **Y** maximizing the probability that an observation from an $N(\mathbf{M}, \mathbf{\Lambda}^{-1})$ distribution lies on the same side of the hyperplane as **M** is (2·3), the tangent at **Y** to the family of constant likelihood surfaces.

Proof. The probability in question is

$$\max\left[\int_{-\infty}^{a} N(0,1)\, dx, \quad \int_{a}^{\infty} N(0,1)\, dx\right],$$

where a is defined by (3·3) with **Y** substituted for $\mathbf{X}_0$. Thus the probability is maximized when $\ln(a^2)$ is maximized with respect to **B**. This is simply obtained by finding

$$\frac{\partial \ln(a^2)}{\partial \mathbf{B}} = \frac{2(\mathbf{M}-\mathbf{Y})}{m} - \frac{2\mathbf{\Lambda}^{-1}\mathbf{B}}{\sigma^2}, \tag{3·4}$$

where m, σ are the numerator and denominator of a. On equating the partial derivative vector (3·4) to **O** we find

$$\mathbf{B} = (\sigma^2/m)\,\mathbf{\Lambda}(\mathbf{M}-\mathbf{Y}). \tag{3·5}$$

As **B** need only be known to within a scale factor the term (σ^2/m) may be omitted from (3·5). This result does in fact represent a maximum; the minimum, $a = 0$, is obscured by the logarithm and corresponds to a hyperplane through **M**. The hyperplane required is therefore: $(\mathbf{X}-\mathbf{Y})'\boldsymbol{\Lambda}(\mathbf{M}-\mathbf{Y}) = 0$, which is equivalent to (2·3).

LEMMA 5. The hyperplane which simultaneously
(i) effects a fixed division of probabilities with respect to one multivariate normal distribution;
(ii) maximizes the division of probabilities with respect to another multivariate normal distribution;
is a fully common tangent to the two families of constant likelihood contours.

Proof. The transformed form (1·2) may be used (with corresponding assumptions) with the distribution of (i) being $N(\mathbf{O}, \mathbf{I})$.

The class of hyperplanes satisfying (i) are the tangents to

$$\mathbf{X}'\mathbf{X} = C, \tag{3·6}$$

where C is a constant whose value depends on the division specified in (i). This follows from the fact that every hyperplane is tangential to one of the family of constant likelihood contours, i.e. can be expressed as

$$(\mathbf{X}-\mathbf{X}_0)'\mathbf{X}_0 = 0 \tag{3·7}$$

for some $\mathbf{X}_0$. This is equivalent to saying that $\mathbf{X}_0$ can be chosen in (3·7) such that this tangent plane divides $N(\mathbf{O}, \mathbf{I})$ into the preassigned probabilities.

The tangent (3·7) divides the space into two regions; the probability that an observation from the $N(\mathbf{M}, \mathbf{D}^{-1})$ distribution will fall on the same side of the tangent as **M** is

$$\max\left[\int_{-\infty}^{b} N(0,1)\,dx, \int_{b}^{\infty} N(0,1)\,dx\right], \tag{3·8}$$

where

$$b = (\mathbf{M}-\mathbf{X}_0)'\mathbf{X}_0/(\mathbf{X}_0'\mathbf{D}^{-1}\mathbf{X}_0)^{\frac{1}{2}}. \tag{3·9}$$

Thus (3·8) is maximized when $\ln(b^2)$ is maximized. To fulfil the conditions of the Lemma this maximum is restricted by the fact that $\mathbf{X}_0$ must satisfy (3·6). By the use of Lagrange multipliers and the regularity of $\ln(b^2)$, maxima, minima, and some points of inflexion of $\ln(b^2)$ occur when

$$\frac{\partial \ln(b^2)}{\partial \mathbf{X}_0} - \frac{\lambda\,\partial(\mathbf{X}_0'\mathbf{X}_0 - C)}{\partial \mathbf{X}_0} = \mathbf{0} \tag{3·10}$$

and

$$\mathbf{X}_0'\mathbf{X}_0 - C = 0. \tag{3·11}$$

(3·10) may be expanded as

$$\frac{2(\mathbf{M}-2\mathbf{X}_0)}{(\mathbf{M}-\mathbf{X}_0)'\mathbf{X}_0} - \frac{2\mathbf{D}^{-1}\mathbf{X}_0}{(\mathbf{X}_0'\mathbf{D}^{-1}\mathbf{X}_0)} - 2\lambda\mathbf{X}_0 = \mathbf{0}. \tag{3·10a}$$

If $\mathbf{X}_0 = \mathbf{Y}(t)$ subject to (3·11) where $\mathbf{Y}(t)$ is defined in (2·8), then

$$(\mathbf{M}-\mathbf{X}_0)'\mathbf{X}_0 = -t\mathbf{M}'(\mathbf{D}-t\mathbf{I})^{-1}\mathbf{D}(\mathbf{D}-t\mathbf{I})^{-1}\mathbf{M} = -t\mathbf{X}_0'\mathbf{D}^{-1}\mathbf{X}_0$$

and

$$\begin{aligned}\mathbf{M}-2\mathbf{X}_0+t\mathbf{D}^{-1}\mathbf{X}_0 &= (\mathbf{I}-2(\mathbf{D}-t\mathbf{I})^{-1}\mathbf{D}+t\mathbf{D}^{-1}(\mathbf{D}-t\mathbf{I})^{-1}\mathbf{D})\,\mathbf{M}\\ &= -(\mathbf{D}-t\mathbf{I})^{-1}\mathbf{D}\mathbf{M} = -\mathbf{X}_0;\end{aligned}$$

thus with

$$\lambda = -t(\mathbf{M}'(\mathbf{D}-t\mathbf{I})^{-1}\mathbf{D}(\mathbf{D}-t\mathbf{I})^{-1}\mathbf{M})^{-1}$$

yield solutions of (3·10); i.e. points of the fully common tangent locus satisfy (3·10). To prove that the maximum is necessarily on this locus, (3·10a) is premultiplied by $\mathbf{X}_0'$ and (3·11) is used to obtain

$$\lambda = -\frac{1}{(\mathbf{M}-\mathbf{X}_0)'\mathbf{X}_0}. \tag{3·12}$$

Using (3·12) in (3·10) and simplifying leads to

$$\frac{(\mathbf{X}-\mathbf{M}_0)}{(\mathbf{M}-\mathbf{X}_0)'\mathbf{X}_0} = \frac{\mathbf{D}^{-1}\mathbf{X}_0}{\mathbf{X}_0'\mathbf{D}^{-1}\mathbf{X}_0}. \tag{3·13}$$

Redefining

$$\mathbf{M}-\mathbf{X}_0 = -t\mathbf{D}^{-1}\mathbf{X}_0, \tag{3·14}$$

where t is the ratio of scalar denominators $t = -(\mathbf{M}-\mathbf{X}_0)'\mathbf{X}_0/\mathbf{X}_0'\mathbf{D}^{-1}\mathbf{X}_0$ which leads to the known relationship (2·8).

As in Lemma 4 there are some minima obscured by the logarithm, which are those tangent to (3·7) passing through **M**; none of these tangents are real when $\mathbf{X}_0'\mathbf{X}_0 > \mathbf{M}'\mathbf{M}$.

Apart from these last tangents we have, in the situation where m_i are non-zero and d_i are distinct, found as many roots as there are dimensions. It has not been verified that there are no other roots.

4. Solution for certain cases

The mathematical description of the cases covered in terms of the standardized problem and of the original parameters will be given at the end of this section. The geometric definition is that the hyperplane corresponding to the optimum discriminant passes between the centres of the populations, i.e. that the optimum discriminator discriminates between the means of the two populations.

The hyperplane corresponding to the discriminator will intersect the locus of fully common tangents at at least one $\mathbf{Y}(t)$, where t is negative (Lemma 2). The hyperplane is the fully common tangent at that point; because of the hyperplanes through $\mathbf{Y}(t)$ it simultaneously minimizes both of the conditional probabilities of misclassification (Lemma 4). Lemma 3 establishes that no paradox will occur. Although this fully common tangent may intersect the locus of fully common tangents at $\mathbf{Y}(t)$ for some positive t, the fully common tangent corresponding to this positive value of t will not pass between the centres, and so one of the conditional probabilities of misclassification would be greater than $\frac{1}{2}$. For the fully common tangent corresponding to the negative t both conditional probabilities of misclassification are less than $\frac{1}{2}$ (Lemma 1).

The determination of the value of t can be achieved only numerically. It should be noted that

(i) $\mathbf{Y}'(t)\,\mathbf{Y}(t) = \sum_i m_i^2 d_i^2/(d_i - t)^2$ is monotonically increasing as t increases from $-\infty$, 0;

(ii) $(\mathbf{Y}(t)-\mathbf{M})'\,\mathbf{D}(\mathbf{Y}(t)-\mathbf{M}) = t^2 \sum_i m_i^2 d_i/(d_i - t)^2$ is monotonically decreasing as t increases from $-\infty$ to 0.

The two weighted conditional probabilities of misclassification are therefore oppositely monotonic in t for t negative. As they are both continuous functions of t the minimax will be achieved when they are equal. Thus numerically the problem becomes that of finding the unique negative root of

$$W_1 \int_{a(t)}^{\infty} N(0,1)\,dx = W_2 \int_{b(t)}^{\infty} N(0,1)\,dx, \tag{4·1}$$

where

$$a(t) = [\sum_i m_i^2 d_i^2/(d_i - t)^2]^{\frac{1}{2}}, \tag{4·1a}$$

$$b(t) = t[\sum_i m_i^2 d_i/(d_i - t)^2]^{\frac{1}{2}}. \tag{4·1b}$$

The condition for the optimum discriminant hyperplane to pass between the centres is the condition for (4·1) to have a negative root. If there is a negative root then this indicates a restricted minimum which, by an argument analogous to that which eliminated the possibility of paradoxes, is an unrestricted minimum.

The condition that there be a negative root is that

$$\left\{\tfrac{1}{2}W_1 - W_2 \int_{b(-\infty)}^{\infty} N(0,1)\,dx\right\}\left\{\tfrac{1}{2}W_2 - W_1 \int_{a(0)}^{\infty} N(0,1)\,dx\right\} > 0, \tag{4·2}$$

where

$$b(-\infty) = [\sum_i m_i^2 d_i]^{\frac{1}{2}}, \tag{4·2a}$$

$$a(0) = [\sum_i m_i^2]^{\frac{1}{2}}. \tag{4·2b}$$

In the original parametrization of $N(\mathbf{M}_1, \mathbf{\Sigma}_1)$, $N(\mathbf{M}_2, \mathbf{\Sigma}_2)$ then

$$b(-\infty) = [(\mathbf{M}_2 - \mathbf{M}_1)'\,\mathbf{\Sigma}_2^{-1}(\mathbf{M}_2 - \mathbf{M}_1)]^{\frac{1}{2}}, \tag{4·3}$$

$$a(0) = [(\mathbf{M}_2 - \mathbf{M}_1)'\,\mathbf{\Sigma}_1^{-1}(\mathbf{M}_2 - \mathbf{M}_1)]^{\frac{1}{2}}. \tag{4·3a}$$

5. Other cases

When the condition (4·2) is not satisfied the problem becomes more complicated. The best discriminating hyperplane will not pass between the centres. Minimax discrimination will 'favour' the population with the greater weight factor. (If the weight factors are equal then (4·2) is satisfied except when $\mathbf{M} = \mathbf{0}$.) From Lemma 5 it follows that the best discriminating hyperplane is still a fully common tangent. The minimization of $W_1P_1 + W_2P_2$ also leads to the family of fully common tangents. Thus both, and indeed any reasonable, criterion for linear discrimination leads, for the multivariate normal populations, to picking one of the family of fully common tangents. The determination of the appropriate one can be achieved only numerically.

Reference

Riffenburgh, R. H. & Clunies-Ross, C. W. (1960). Linear discriminant analysis. *Pacif. Sci.* **14**, No. 3, July 1960 (In Press).

Approximation to the distribution of the sample size for sequential tests. II. Tests of composite hypotheses

By D. H. BHATE

University College London and Delhi School of Economics, Delhi

1·1. *Introduction.* The original method of construction of sequential probability ratio tests given by Wald (1947) applied to discrimination between two simple hypotheses. Wald extended his procedure to tests discriminating between composite hypotheses by introducing a rather arbitrary weighting function, by means of which the problem was reduced to choice between simple hypotheses. Armitage (1947), Barnard (1952) and Cox (1952) gave, in different ways, direct probability justification for using appropriate likelihood ratios for comparisons of certain composite hypotheses. Nandi (1946) and Bahadur (1954) have also discsussed this problem. Rushton (1950, 1952*a*, *b*) studied sequential versions of the t and F tests. Johnson (1953, 1954) considered sequential tests in general analysis of variance problems. Ray (1956) gives a detailed discussion of the application of these tests in one-way classifications by groups and in randomized blocks.

In general, we do not know much about the distribution of sample size, or even indeed the average sample size (a.s.n.) for these sequential tests. In this paper the methods used by Bhate (1959) in studying the distribution of the decisive sample size (d.s.n.) for sequential tests of simple hypotheses will be extended to tests of composite hypotheses.

2·1. Let us suppose that Π_1 and Π_2 are two Normal populations with means μ_1, μ_2 and variances σ_1^2, σ_2^2, respectively, with $\lambda = \sigma_1^2/\sigma_2^2$. We wish to discriminate between the (composite) hypotheses H_0: $\lambda = \lambda_0$ and H_1: $\lambda = \lambda_1$. If we wish to control (approximately) the chances of error by making

$$\Pr\{\text{accept } H_1 \mid H_0\} \doteqdot \alpha, \quad \Pr\{\text{accept } H_0 \mid H_1\} \doteqdot \beta$$

an appropriate sequential test is to measure one individual from each of the two populations at each stage, giving measurements x_{1m}, x_{2m} from Π_1, Π_2, respectively, at the mth stage, calculate

$$F_m = \sum_{i=1}^{m} (x_{1i}-\bar{x}_{1m})^2 \Big/ \sum_{i=1}^{m} (x_{2i}-\bar{x}_{2m})^2, \quad \text{where} \quad \bar{x}_{tm} = \frac{1}{m}\sum_{i=1}^{m} x_{ti} \quad (t = 1, 2)$$

and (i) continue to the $(m+1)$th stage if $F_m^- < F_m < F_m^+$,
(ii) accept H_0 if $F_m^- \geqslant F_m$,
(iii) accept H_1 if $F_m^+ \leqslant F_m$,
where F_m^-, F_m^+ are, respectively, solutions of the equations

$$\frac{\beta}{1-\alpha}, \quad \frac{1-\beta}{\alpha} = \left(\frac{\lambda_0}{\lambda_1}\right)^{\frac{1}{2}(m-1)} \left\{\frac{1+F_m/\lambda_0}{1+F_m/\lambda_1}\right\}^{m-1}.$$

Upper and lower bounds for the cumulative distribution of the d.s.n. (Bhate, 1959, equation (2·2·4)) are given by the inequalities

$$P_m \leqslant \sum_{i=1}^{m} p_i \leqslant P_m + \sum_{i=1}^{m-1} p_{iA}\,\eta_{im} + \sum_{i=1}^{m-1} p_{iR}\,\delta_{im}, \tag{2·1·1}$$

where (*a*) p_i = probability that a decision is reached at (but not before) the ith stage;

$\left(\text{i.e. } \sum_{i=1}^{m} p_i = \text{probability that the d.s.n. is less than or equal to } m\right)$;

(*b*) $P_m = \int_0^{F_m^-} p(F_m)\,dF_m + \int_{F_m^+}^{\infty} p(F_m)\,dF_m$;

(*c*) $p_{iA}(p_{iR})$ = probability of deciding to accept $H_0(H_1)$ at the ith stage;

and (*d*) $\eta_{im}(\delta_{im})$ = the largest possible value for the conditional probability that, a decision having been reached at the ith stage, $(m-i)$ further observations would produce a value of F_m such that $F_m^- \leqslant F_m \leqslant F_m^+$.

The quantities F_m^- and F_m^+ are fixed numbers and can be calculated directly. The lower bound of the cumulative distribution of the d.s.n., namely

$$P_m = \int_0^{F_m^-} p(F_m \mid \lambda)\,dF_m + \int_{F_m^+}^{\infty} p(F_m \mid \lambda)\,dF_m$$

is a function of the true value of $\lambda = \sigma_1^2/\sigma_2^2$, and is easily calculated.

The calculation of the upper bound is more complicated. It calls for splitting up p_i into p_{iA} and p_{iR}, and evaluating η_{im} and δ_{im}.

Under certain conditions of symmetry, and supposing $\alpha = \beta$, we have p_{iA}/p_{iR} constant with respect to i, so that when H_0 is true

$$p_{iA} = (1-\alpha)\,p_i, \quad p_{iR} = \alpha p_i$$

and when H_1 is true

$$p_{iA} = \alpha p_i, \quad p_{iR} = (1-\alpha)\,p_i.$$

Calculation of η_{im}.

$$\eta_{im} = \max_{F_i \leqslant F_i^-} \int_{F_m^-}^{F_m^+} p(F_m \mid F_i)\,dF_m.$$

It is evident that the maximum value of the integral is attained when $F_i = F_i^-$. Hence we must evaluate

$$\eta_{im} = \int_{F_m^-}^{F_m^+} p(F_m \mid F_i^-)\,dF_m.$$

Now $F_m = \lambda\chi^2_{1,m-1}/\chi^2_{2,m-1}$, where the numerator and denominator contain independent χ^2's, each with $m-1$ degrees of freedom. With an obvious notation, we can write

$$F_m = \lambda\frac{\chi^2_{1,i-1}+\chi^2_{1,m-i}}{\chi^2_{2,i-1}+\chi^2_{2,m-i}},$$

all the χ^2's being mutually independent. Given that $\lambda\chi^2_{1,i-1}/\chi^2_{2,i-1} = F_i^-$, it is straightforward to show that the conditional distribution of $\chi^2_{2,i-1}$ is that of $(1+F_i^-/\lambda)^{-1}$ times a χ^2-variable with $2(i-1)$ degrees of freedom. We will denote this by $(1+F_i^-/\lambda)^{-1}\chi^2_{2(i-1)}$. This quantity is of course independent of $\chi^2_{1,m-i}$ and $\chi^2_{2,m-i}$. Hence

$$\Pr\{F_m \leqslant F_m^- \mid F_i = F_i^-\} = \Pr\left\{\frac{F_i^-/\lambda}{1+F_i^-/\lambda}\chi^2_{2(i-1)} + \chi^2_{1,(m-i)} \leqslant \lambda^{-1}F_m^-\left(\frac{1}{1+F_i^-/\lambda}\chi^2_{2(i-1)} + \chi^2_{2,(m-i)}\right)\right\} \tag{2·1·2}$$

and so, if $\lambda = 1$,

$$\Pr\{F_m \leqslant F_m^- \mid F_i = F_i^-\} = \Pr\left\{\frac{\chi^2_{1,m-i}}{\left(\frac{F_m^- - F_i^-}{1+F_i^-}\right)\chi^2_{2(i-1)} + F_m^-\chi^2_{2,(m-i)}} \leqslant 1\right\}. \tag{2·1·3}$$

Using the approximation

$$\lambda_1\chi^2_{\nu_1} + \lambda_2\chi^2_{\nu_2} \doteqdot g\chi^2_\nu$$

with $$g = (\lambda_1^2\nu_1 + \lambda_2^2\nu_2)/(\lambda_1\nu_1 + \lambda_2\nu_2), \quad \nu = (\lambda_1\nu_1 + \lambda_2\nu_2)^2/(\lambda_1^2\nu_1 + \lambda_2^2\nu_2)$$

(see, for example, Welch, 1938; Box, 1954), we have

$$\Pr\{F_m \leqslant F_m^- \mid F_i = F_i^-\} \doteqdot \Pr\left\{\frac{\chi^2_{1,m-i}}{g\chi^2_\nu} \leqslant 1\right\}, \tag{2·1·4}$$

where

$$g = \frac{\left(\frac{F_m^- - F_i^-}{1+F_i^-}\right)^2 2(i-1) + (F_m^-)^2(m-i)}{\left(\frac{F_m^- - F_i^-}{1+F_i^-}\right)2(i-1) + F_m^-(m-i)},$$

$$\nu = \frac{\left\{\left(\frac{F_m^- - F_i^-}{1+F_i^-}\right)2(i-1) + F_m^-(m-i)\right\}^2}{\{(F_m^- - F_i^-)/(1+F_i^-)\}^2\,2(i-1) + (F_m^-)^2(m-i)}.$$

The right-hand side of (2·1·4) can be evaluated as the incomplete beta-function $I_{g/(1+g)}(\frac{1}{2}(m-i), \frac{1}{2}\nu)$ and so, an approximate value for η_{im} can be obtained.

2·2. We will now consider two illustrative examples.

Example 1

Let H_0 specify $\lambda = \lambda_0 = 1$, H_1 specify $\lambda = \lambda_1 = 16$

with $\alpha = \beta = 0{\cdot}05$.

The continuation region after the mth stage is then

$$\tfrac{1}{19} < (\tfrac{1}{16})^{\frac{1}{2}(m-1)}\left(\frac{1+F_m}{1+F_m/16}\right)^{m-1} < 19$$

The boundary values for F_m are shown in Table 1.

We will consider the distribution of the d.s.n. when H_0 is true, i.e. when $\lambda = 1$.

In the present case it is possible to compute an approximate value for η_{im} (and similarly for δ_{im}). Hence upper and lower bounds for the cumulative distribution function of the d.s.n. can be computed. We will take the mean of these bounds to get a single estimate of the cumulative distribution function of the d.s.n. As a check on these estimates, the distribution was also computed by quadrature, forming the conditional distributions of F_m, stage by stage.

The results of these calculations are shown in Table 2 where the figures in the fourth column are the differences of $\frac{1}{2}(L_m + U_m)$.

Table 1

m	F_m^-	F_m^+	m	F_m^-	F_m^+
4	0·550	29·303	8	1·941	8·174
5	1·041	15·393	9	2·135	7·475
6	1·415	11·346	10	2·300	6·937
7	1·710	9·309			

Table 2

	Cumulative Σp_m			
m	Lower bound L_m	Upper bound U_m	Approximate p_m	p_m by quadrature
4	0·32066	0·32066	0·32066	0·32066
5	·52534	·58840	·23621	·23793
6	·65933	·77988	·16274	·14465
7	·74155	·89241	·09738	·09275
8	·80598	·95524	·06365	·06835
9	·85193	·98126	·03598	·04297

Example 2

Let H_0 specify $\lambda = \lambda_0 = 1$, H_1 specify $\lambda = \lambda_1 = 17$

with $\alpha = \beta = 0{\cdot}01$. The continuation region is then

$$\tfrac{1}{99} < (\tfrac{1}{17})^{\frac{1}{2}(m-1)} \left(\frac{1+F_m}{1+F_m/17} \right)^{m-1} < 99.$$

The approximate distribution of the d.s.n. and its upper and lower bounds, calculated by the methods used in Example 1 are given in Table 3. The figures in the fourth column have been calculated as for Table 2.

Table 3

	Cumulative distribution		
m	Lower bound L_m	Upper bound U_m	Approximate p_m
5	0·15743	0·15743	0·15743
6	·36758	·38821	·22046
7	·52071	·56987	·16740
8	·62224	·70606	·11886
9	·71138	·79579	·08943
10	·78089	·87354	·07363
11	·82917	·91806	·04640
12	·87299	·95510	·04043

Quadrature was not applied in this Example; the accuracy of the approximations should be similar to that obtained in Example 1.

2·3. In many cases it is not possible to estimate η_{im}, even approximately, by the methods outlined in this paper. For example, this cannot be done for the sequential procedure in a one-way classification (by groups), with a fixed number of groups, described by Johnson (1954). (The calculation is possible, on the other hand, for a sequential procedure in which the number of groups is increased in the course

of sampling). However, even when the methods used in Examples 1 and 2 are not applicable, it is always possible to evaluate the lower bound, P_m, and upper bounds can often be found by *ad hoc* methods (as in Example 3 below).

2·4. *Sequential t-tests.* For the sequential t-test, discriminating between two values of the mean of a Normal population with unknown variance, described by Rushton (1952), it is not possible to calculate η_{im}. However, an upper bound is provided by the cumulative distribution function of the d.s.n. for the standard sequential procedure discriminating between the same two values of the mean when the population variance is known.

Example 3

Let us suppose that x is a Normal variable with mean θ and standard deviation σ; write $\delta = (\theta - \theta_0)/\sigma$. We want to discriminate between

$$H_0 \text{ specifying } \delta = \delta_0 = 0 \quad \text{and} \quad H_1 \text{ specifying } \delta = \delta_1 = 1, \quad \text{with} \quad \alpha = \beta = 0{\cdot}05.$$

This can be done by a one-sided sequential t-test with boundary values given by Rushton (1952). Table 4 gives lower and upper bounds for the distribution of the d.s.n. calculated as described above.

Table 4

	Cumulative distribution		
m	Lower bound L_m	Upper bound U_m	Approximate p_m
4	0·33217	0·36117	0·34667
6	·52210	·57979	·20432
8	·64220	·72315	·13172
10	·73179	·82163	·09400
15	·86414	·94316	·12694
20	·93066	·99008	·05672
25	·96229	1·00000	·02077

The approximate values of p_m in the last column have been calculated in the same way as in the preceding examples.

The author would like to thank Dr N. L. Johnson for his helpful suggestions during the investigation of this problem.

References

Armitage, P. (1947). Some sequential tests of Student's hypotheses. *J. R. Statist. Soc.* Suppl. **9**, 250–63.

Bahadur, R. R. (1954). Sufficiency and statistical decision functions. *Ann. Math. Statist.* **25**, 423–63.

Barnard, G. A. (1952). The frequency justification of certain sequential tests. *Biometrika*, **39**, 144–50.

Bhate, D. H. (1959). Approximation to the distribution of the sample size for sequential tests. I. Tests of simple hypotheses. *Biometrika*, **46**, 130–8.

Box, G. E. P. (1954). Some theorems on quadratic forms. *Ann. Math. Statist.* **25**, 290–302.

Cox, D. R. (1952). Sequential tests of composite hypotheses. *Proc. Camb. Phil. Soc.* **48**, 290–9.

Johnson, N. L. (1953). Some notes on the application of sequential methods to the analysis of variance. *Ann. Math. Statist.* **24**, 614–23.

Johnson, N. L. (1954). Sequential procedure in certain component of variance problems. *Ann. Math. Statist.* **25**, 357–66.

Nandi, H. K. (1946). Use of well known statistics in sequential analysis. *Sankhyā*, **8**, 339–44.

Ray, W. D. (1956). Sequential analysis applied to certain experimental designs in the analysis of variance. *Biometrika*, **43**, 388–403.

Rushton, S. (1950). On a sequential t-test. *Biometrika*, **37**, 326–33.

Rushton, S. (1952). Ph.D. Thesis, University of London.

Rushton, S. (1952*a*). On a two-sided sequential t-test. *Biometrika*, **39**, 302–8.

Rushton, S. (1952*b*). Sequential tests of equality of variance of two normal populations, with known means. *Sankhyā*, **12**, 63–78.

Wald, A. (1947). *Sequential Analysis.* New York: John Wiley and Sons.

Welch, B. L. (1938). Significance of the difference between two means when the population variances are unequal. *Biometrika*, **29**, 350–62.

On the covariance determinants of moving-average and autoregressive models

BY P. D. FINCH

Research Techniques Division, London School of Economics

1. INTRODUCTION

It has been noted (Durbin, 1959) that the limiting covariance determinant of n successive values of the second-order moving-average series (1)

$$x_t = \epsilon_t + \beta_1 \epsilon_{t-1} + \beta_2 \epsilon_{t-2}, \tag{1}$$

where the ϵ_t have zero mean, unit variance and are uncorrelated is the same as the limiting covariance determinant of n successive values of the autoregressive series (2)

$$x_t + \beta_1 x_{t-1} + \beta_2 x_{t-2} = \epsilon_t \tag{2}$$

with same coefficients and the same residual variance.

In this note we show that this result is generally true. We formulate the generalization for discrete weakly stationary processes with rational spectral density functions although it will be seen that the proof will apply to processes with absolutely continuous spectral distribution functions whose spectral densities are reciprocals of each other and satisfy a Lipshitz condition (see formulation of Theorem *GS* below). Explicitly we prove the following theorem.

THEOREM 1. Let

$$A(z) = a_0 + a_1 z + a_2 z^2 + \ldots + a_p z^p \quad (a_0 = 1),$$
$$B(z) = b_0 + b_1 z + b_2 z^2 + \ldots + b_q z^q \quad (b_0 = 1),$$

where p, q are integral, the coefficients a_i, b_j are real and all the roots of $A(z)$ and all those of $B(z)$ lie without the unit circle $|z| = 1$. Let $\{\epsilon_i\}$, $\{\eta_i\}$ be two independent (real) processes with

$$E\{\epsilon_t\} = E\{\eta_t\} = 0, \quad E\{\epsilon_t \epsilon_s\} = E\{\eta_t \eta_s\} = \delta_{s,t}.$$

Consider the two weakly stationary (real) processes $\{x_t\}$, $\{y_t\}$ satisfying respectively the equations

$$a_0 x_t + a_1 x_{t-1} + \ldots + a_p x_{t-p} = b_0 \epsilon_t + b_1 \epsilon_{t-1} + \ldots + b_q \epsilon_{t-q}, \tag{3}$$
$$a_0 \eta_t + a_1 \eta_{t-1} + \ldots + a_p \eta_{t-p} = b_0 y_t + b_1 y_{t-1} + \ldots + b_q y_{t-q}. \tag{4}$$

Let X_n, Y_n be the covariance determinants of order $(n+1)$ of the processes $\{x_t\}$, $\{y_t\}$. That is

$$X_n = \det \|E\{x_s x_t\}\| \quad (s, t = 0, 1, \ldots, n),$$
$$Y_n = \det \|E\{y_s y_t\}\| \quad (s, t = 0, 1, \ldots, n).$$

Then

$$\lim_{n \to \infty} X_n = \lim_{n \to \infty} Y_n. \tag{5}$$

Further the common limiting value of these determinants is given explicitly by the following expression

$$\left| \frac{\beta_1 \beta_2 \ldots \beta_q}{\alpha_1 \alpha_2 \ldots \alpha_p} \right|^{2(q-p)} \frac{\Pi |\beta_\mu \overline{\alpha}_\nu - 1|^2}{\Pi |\beta_\mu \overline{\beta}_{\mu'} - 1| \, \Pi |\alpha_\nu \overline{\alpha}_{\nu'} - 1|} \quad \begin{matrix} (\mu, \mu' = 1, 2, \ldots, q), \\ (\nu, \nu' = 1, 2, \ldots, p), \end{matrix} \tag{6}$$

where $\alpha_1, \alpha_2, \ldots, \alpha_p$ are the distinct roots of $A(z)$ and $\beta_1, \beta_2 \ldots \beta_q$ are the distinct roots of $B(z)$.

2. PROOF OF THE THEOREM

The theorem may be deduced very simply from a general theorem on the determinants of Toeplitz forms. The general theorem is proved in Grenander & Szego (1958, § 5·5) and for our purpose may be formulated as follows.

THEOREM *GS*. Let $f(\lambda)$ be positive and continuous in $[-\pi, \pi]$ with $f(-\pi) = f(\pi)$. Let $f'(\lambda)$ exist and satisfy the Lipshitz condition

$$|f'(\lambda_1) - f'(\lambda_2)| < K |\lambda_1 - \lambda_2|^\alpha \quad (K > 0, \, 0 < \alpha < 1).$$

Let

$$T_n(f) = \frac{1}{2\pi} \int_{-\pi}^{+\pi} |u_0 + u_1 e^{i\lambda} + \ldots + u_n e^{in\lambda}|^2 f(\lambda) \, d\lambda$$

be the nth Toeplitz form associated with $f(\lambda)$ and let $D_n(f)$ be the determinant of this form. Then

$$\lim_{n\to\infty}\frac{D_n(f)}{\{G(f)\}^{n+1}}=\exp\left\{\frac{1}{\pi}\iint\left|\frac{g'(z)}{g(z)}\right|^2 d\sigma\right\}, \tag{7}$$

where $G(f)=g(0)$ and $g(z)$ is a function uniquely determined by $f(\lambda)$ satisfying

$$[g(0)]^2=\exp\left\{\frac{1}{2\pi}\int_{-\pi}^{+\pi}\log f(\lambda)\,d\lambda\right\},$$

$$|g(e^{i\lambda})|^{-2}=f(\lambda)$$

and the integration on the right of (7) extends over the unit circle $|z|\leqslant 1$.

It is well known that the processes $\{x_t\}$, $\{y_t\}$ defined by (3) and (4) have absolutely continuous spectral distribution functions with spectral densities $(2\pi)^{-1}f_x(\lambda)$, $(2\pi)^{-1}f_y(\lambda)$ given by

$$f_x(\lambda)=\left|\frac{B(e^{i\lambda})}{A(e^{i\lambda})}\right|^2,$$

$$f_y(\lambda)=\left|\frac{A(e^{i\lambda})}{B(e^{i\lambda})}\right|^2.$$

The covariances of the $\{x_t\}$, $\{y_t\}$ processes are given by

$$E\{x_s x_t\}=\frac{1}{2\pi}\int_{-\pi}^{+\pi}e^{i(s-t)\lambda}f_x(\lambda)\,d\lambda,$$

$$E\{y_s y_t\}=\frac{1}{2\pi}\int_{-\pi}^{+\pi}e^{i(s-t)\lambda}f_y(\lambda)\,d\lambda$$

so the covariance determinants X_n, Y_n are just the determinants of the Toeplitz forms $T_n(f_x)$, $T_n(f_y)$, respectively.

To apply Theorem *GS* we remark that the functions $g_x(z)$, $g_y(z)$ are given by

$$g_x(z)=\frac{A(z)}{B(z)},$$

$$g_y(z)=\frac{B(z)}{A(z)}$$

and that $G(f_x)=g_x(0)=G(f_y)=g_y(0)=1$. Thus we have

$$\lim_{n\to\infty}X_n=\exp\left\{\frac{1}{\pi}\iint\left|\frac{g_x'(z)}{g_x(z)}\right|^2 d\sigma\right\},$$

$$\lim_{n\to\infty}Y_n=\exp\left\{\frac{1}{\pi}\iint\left|\frac{g_y'(z)}{g_y(z)}\right|^2 d\sigma\right\}.$$

But $$g_x(z)=[g_y(z)]^{-1} \quad \text{hence} \quad \frac{g_x'(z)}{g_x(z)}=-\frac{g_y'(z)}{g_y(z)}$$

and consequently $\lim_{n\to\infty}X_n=\lim_{n\to\infty}Y_n$ which proves (5).

In order to obtain (6) we evaluate the integral on the right of (7). Note that

$$\frac{g_x'(z)}{g_x(z)}=\frac{A'(z)}{A(z)}-\frac{B'(z)}{B(z)}.$$

Then if $A(z)=A(\alpha_1-z)(\alpha_2-z)\dots(\alpha_p-z)$ and $B(z)=B(\beta_1-z)(\beta_2-x)\dots(\beta_q-z)$ we obtain

$$\frac{g_x'(z)}{g_x(z)}=\sum_{\mu=1}^{q}\frac{1}{\beta_\mu-z}-\sum_{\nu=1}^{p}\frac{1}{\alpha_\nu-z}$$

and

$$\left|\frac{g_x'(z)}{g_x(z)}\right|^2=\sum_{\mu,\mu'=1\dots q}\frac{1}{(\beta_\mu-z)(\overline{\beta}_{\mu'}-\overline{z})}+\sum_{\nu,\nu'=1\dots p}\frac{1}{(\alpha_\nu-z)(\overline{\alpha}_{\nu'}-\overline{z})}$$
$$-\sum_{\substack{\mu=1\dots q\\ \nu'=1\dots p}}\frac{1}{(\beta_\mu-z)(\overline{\alpha}_{\nu'}-\overline{z})}-\sum_{\substack{\mu'=1\dots q\\ \nu=1\dots p}}\frac{1}{(\overline{\beta}_{\mu'}-\overline{z})(\alpha_\nu-z)}.$$

Noting that

$$\frac{1}{\pi}\iint\frac{d\sigma}{(z-k)(\overline{z}-l)}=\log\left\{\frac{lk}{lk-1}\right\}$$

we obtain

$$\frac{1}{\pi}\iint\left|\frac{g_x'(z)}{g_x(z)}\right|^2 dz=\log\left[\frac{\Pi(\beta_\mu\overline{\beta}_{\mu'})\,\Pi(\alpha_\nu\overline{\alpha}_{\nu'})}{\Pi(\beta_\mu\overline{\beta}_{\mu'}-1)\,\Pi(\alpha_\nu\overline{\alpha}_{\nu'}-1)}\Big/\frac{\Pi(\beta_\mu\overline{\alpha}_{\nu'})\,\Pi(\overline{\beta}_{\mu'}\alpha_\nu)}{\Pi(\beta_\mu\overline{\alpha}_{\nu'}-1)\,\Pi(\overline{\beta}_{\mu'}\alpha_\nu-1)}\right],$$

whence

$$\lim_{n\to\infty} X_n = \left|\frac{\beta_1\beta_2\dots\beta_q}{\alpha_1\alpha_2\dots\alpha_p}\right|^{2(q-p)} \frac{\Pi\,|\beta_\mu\bar{\alpha}_\nu - 1|^2}{\Pi\,|\beta_\mu\bar{\beta}_{\mu'} - 1|\,\Pi\,|\alpha_\nu\bar{\alpha}_{\nu'} - 1|} \quad (\mu,\mu' = 1,2,\dots,q;\ \nu,\nu' = 1,2,\dots,p).$$

This proves (6).

This result is easily modified if $A(z)$ or $B(z)$ have multiple roots.

This paper was written under a grant from the Ford Foundation.

References

Durbin, J. (1959). Efficient estimation of parameters in moving-average models. *Biometrika*, **46**, 306–16.

Grenander, U. & Szego, G. (1958). *Toeplitz Forms and their Application*. University of California Press.

A note on some approximations to the variance in discrete-time stochastic models for biological systems

By P. H. LESLIE

Bureau of Animal Population, Department of Zoological Field Studies, Oxford

In the original description of a discrete-time stochastic model for studying the properties of certain biological systems by numerical methods, (Leslie, 1958), certain empirical approximations to the variance $\sigma^2(N_{t+1}\,|\,N_t)$ were suggested in order to save machine time at each step in the calculations. Although these approximations appear to work very well in practice (cf. some of the results given in Bartlett, Gower & Leslie, 1960), their scope is very limited, and it therefore may be a convenience if some more general approximations are available, which should hold over a wider range of possible values for the constant birth-rate (b) and the constant death-rate (d).

Considering only one of the species in a system, and given that this population consists of N_t individuals at time t, then the expected number of individuals at time $t+1$ is

$$E\{N_{t+1}\,|\,N_t\} = \frac{\lambda}{q_t}N_t = \lambda_t N_t, \qquad (1)$$

and the variance

$$\sigma^2\{N_{t+1}\,|\,N_t\} = \phi_t E\{N_{t+1}\,|\,N_t\},$$

in which $q_t > 1$ is some function of the numbers at time t in each of the interacting populations, and the constant

$$\lambda = e^{b-d} \quad (b > d). \qquad (2)$$

When the birth-rate (b) of the particular species is assumed to remain constant (BRC model), we have in (1)

$$\left.\begin{aligned} \log_e\lambda_t = r_t &= b - d_t, \\ \phi_t = \left(\frac{2b}{r_t} - 1\right)(\lambda_t - 1) &\quad (\lambda_t \neq 1), \\ = 2b &\quad (\lambda_t = 1), \end{aligned}\right\} \qquad (3)$$

and when the death-rate (d) remains constant, (DRC model),

$$\left.\begin{aligned} \log_e\lambda_t = r_t &= b_t - d \quad (\lambda \geqslant \lambda_t \geqslant e^{-d}). \\ \phi = \left(\frac{2d}{r_t} + 1\right)(\lambda_t - 1) &\quad (\lambda_t \neq 1,\ \lambda \geqslant \lambda_t \geqslant e^{-d}), \\ = 2d &\quad (\lambda_t = 1). \end{aligned}\right\} \qquad (4)$$

In designing a numerical experiment with these models, some restrictions must be placed on the values of b and λ which may be adopted in (2), because of the difference between discrete-time and continuous-time models in regard to the degree of variability in numbers when a system is fluctuating in the region of the stationary state. The variances and covariances of a discrete-time model are always greater than those of its continuous-time equivalent (a point which was not appreciated in Leslie, 1958), but by reducing the unit of time adopted in the discrete-time model, it is always possible to make the

variances of the distributions for the two types of model more in agreement (Bartlett *et al.* 1960). A convenient set of restrictions in the choice of the parameters in (2) has been found to be

$$\lambda = e^{b-d} \leqslant 2{\cdot}0 \quad (b \leqslant 1{\cdot}0). \tag{5}$$

If, in the expression for the variance in (1), we regard ϕ_t as a function of λ_t, we have, for example, from (3),

$$\frac{d\phi_t}{d\lambda_t} = \{2b(e^{-r_t} + r_t - 1)/r_t^2\} - 1,$$

or

$$\frac{d\phi_t}{d\lambda_t} = b\left(1 - \frac{r_t}{3} + \frac{r_t^2}{12} - \ldots\right) - 1.$$

Neglecting the terms in r_t, r_t^2, etc., then

$$\frac{d\phi_t}{d\lambda_t} \sim (b-1), \quad \text{constant},$$

and hence, to a first approximation, we have for the BRC model,

$$\phi_t \sim (b+1) + (b-1)\lambda_t, \tag{6}$$

since from (3) $\phi_t = 2b$, when $\lambda_t = 1$.

Similarly, it may be shown from (4) that as a first approximation for the DRC model,

$$\phi_t \sim (d-1) + (d+1)\lambda_t. \tag{7}$$

As an example of these approximations, consider three values of constant b, viz. 1·0, 0·8 and 0·5. The first two of these represent the more extreme values of b which might be adopted in the BRC model, and are compatible with assumed values of $\lambda \leqslant 2{\cdot}0$; while in the case of constant $b = 0{\cdot}5$, any value of the constant λ must necessarily be less than $\lambda \sim 1{\cdot}65$. The following are the exact and approximate values of ϕ_t over a range of values of λ_t.

	Exact ϕ_t			Approximate ϕ_t		
λ_t	$b = 1{\cdot}0$	$b = 0{\cdot}8$	$b = 0{\cdot}5$	$b = 1{\cdot}0$	$b = 0{\cdot}8$	$b = 0{\cdot}5$
1·9	1·904	1·344	—	2·0	1·42	—
1·7	1·938	1·411	—	2·0	1·46	—
1·5	1·966	1·473	0·733	2·0	1·50	0·75
1·3	1·987	1·530	0·843	2·0	1·54	0·85
1·1	1·998	1·579	0·949	2·0	1·58	0·95
1·0	2·000	1·600	1·000	2·0	1·60	1·00
0·9	1·998	1·618	1·049	2·0	1·62	1·05
0·8	1·993	1·634	1·096	2·0	1·64	1·10
0·7	1·982	1·646	1·141	2·0	1·66	1·15
0·6	1·966	1·653	1·183	2·0	1·68	1·20
0·5	1·943	1·654	—	2·0	1·70	—

It will be seen that these approximations are just as good as those originally suggested in Leslie (1958), while for smaller values of the constant birth-rate (b) they become even better. Similarly, for assumed values of the constant death-rate (d) less than $d = 0{\cdot}6$, say, it was found that the use of (7) gave a good approximation to the exact values of ϕ_t for variable λ_t, the agreement again becoming better as the values of d became smaller.

These approximations for the BRC and DRC models originally arose in considering the behaviour of $\sigma^2\{N_{t+1} \mid N_t\}$ for a logistic process in the neighbourhood of the stationary state (Bartlett *et al.* 1960, §§ 3 and 11); but it appears now that they will also hold over the entire development of any process, provided the restrictions made in (5) in regard to the parameters λ and b are observed. Their use results in a considerable saving of machine time at each step in the calculations for two or more interacting species, since the necessity of including a log subroutine in the programmes is avoided.

References

Bartlett, M. S., Gower, J. C. & Leslie, P. H. (1960). A comparison of theoretical and empirical results for some stochastic population models. *Biometrika*, **47**, 1–11.

Leslie, P. H. (1958). A stochastic model for studying the properties of certain biological systems by numerical methods. *Biometrika*, **45**, 16–31.

On two queues in parallel

BY COLERIDGE A. WILKINS

University of New South Wales

In a paper published in this journal, Haight (1958) investigated a two-queue system in which an arrival is assigned to the shorter queue, or if they are of equal length, to a particular one, the 'near' queue.

The purpose of this note is to show that Dr Haight's results can be extended to the more general case where, if X is the length of the near queue, Y that of the other, and $W(X, Y)$ is the probability of an arrival's joining the near one, we have

$$W(x,y) = \begin{cases} 1 & x < y, \\ w(x) & x = y, \\ 0 & x > y. \end{cases}$$

The new assumption does not affect the expressions for $A(s,s')$, $B(s,s')$, or $C(s,s')$, but does change the values of the functions $I(x,y)$ and $J(x,y)$. The modified values of the latter two are tabled below

$I(x,y)$	$J(x,y)$	
1	0	$x < y-1$
1	$1-w(x) = \chi(x)$	$x = y-1$
1	1	$x = y$
$w(x)$	1	$x = y+1$
0	1	$x > y+1$

We define M_x by $M_x = \chi(x)p_{xx}$.

$D(s,s')$ is now given by

$$D(s,s') = \lambda\left\{\sum_{y=1}^{\infty}\sum_{x=1}^{y} p_{x-1,y}s^x s'^y + \sum_{0}^{\infty}(1-M_y)s^{y+1}s'^y + \sum_{x=1}^{\infty}\sum_{y=1}^{x} p_{x,y-1}s^x s'^y + \sum_{0}^{\infty} M_x s^x s'^{x+1}\right\}.$$

We denote this more general form of D by G, and its value when $w(x)$ is constant and equal to one, by H. If we formally subtract, we find

$$G - H = \lambda\sum_{0}^{\infty} M_k s^k s'^k(s'-s) = \lambda F, \quad \text{say.}$$

The generating function equation for the general case is therefore

$$(1-s')\phi + (s'-s)\psi = \frac{1}{\rho}\frac{1-s}{s}[\phi - pg_0(s')] + \frac{1}{\rho'}\frac{1-s'}{s'}[\phi - p'g_0(s)] + F.$$

The values of F and its first and second derivatives at the points considered by Haight follow.

	$s = s' = 0$	$s = 1, s' = 0$	$s = 0, s' = 1$	$s = s' = 1$
F	0	$-M_0$	M_0	0
F_s	$-M_0$	$-M_0$	$M_1 - M_0$	$-\sum_{0}^{\infty} M_k$
$F_{s'}$	M_0	$M_0 - M_1$	M_0	$\sum_{0}^{\infty} M_k$
F_{ss}	0	0	$2(M_2 - M_1)$	$-2\sum_{0}^{\infty} kM_k$
$F_{s's'}$	0	$2(M_1 - M_2)$	0	$2\sum_{0}^{\infty} kM_k$
$F_{ss'}$	0	$-2M_1$	$2M_1$	0

Haight's equations (16)–(25), (28), (29), (31), (33), (35) and (36) become

$$-\lambda p' + \mu' p_{.1} + \lambda p_{00} w(0) = 0, \tag{16}$$

$$-\lambda p + \mu p_{1.} + \lambda p_{00} \chi(0) = 0, \tag{17}$$

$$-\lambda p_{10} - \mu p_{10} + \mu' p_{11} + \mu p_{20} + \lambda p_{00} w(0) = 0, \tag{18}$$

$$-\lambda p' m_0 + \mu' p_{.1} m_1 - p' \mu + \mu p_{00} + \lambda p_{00} w(0) = 0, \tag{19}$$

$$-\lambda p_{1.} - \mu p_{1.} + \mu p_{2.} + \lambda p + \lambda p_{10} + \lambda(\chi(1) p_{11} - \chi(0) p_{00}) = 0, \tag{20}$$

$$\lambda N = \mu(1-p) + \lambda \sum_0^\infty \chi(k) p_{kk}, \tag{21}$$

$$-\lambda p_{01} - \mu' p_{01} + \mu' p_{02} + \mu p_{11} + \lambda \chi(0) p_{00} = 0, \tag{22}$$

$$-\lambda p_{.1} - \mu' p_{.1} + \mu' p_{.2} + \lambda p' - \lambda w(0) p_{00} + \lambda p_{01} + \lambda w(1) p_{11} = 0, \tag{23}$$

$$-\lambda p m_0' - \mu' p + \mu' p_{00} + \mu p_{1.} m_1' + \lambda \chi(0) p_{00} = 0, \tag{24}$$

$$\lambda F = \mu'(1-p') - \lambda \sum_0^\infty \chi(k) p_{kk}, \tag{25}$$

$$-\lambda p_{2.} - \mu p_{2.} + \mu p_{3.} + \lambda p_{1.} - \lambda p_{10} + \lambda p_{20} + \lambda p_{21} + \lambda(\chi(2) p_{22} - \chi(1) p_{11}) = 0, \tag{28}$$

$$D_{ss}(1,1) = \lambda w + 2\mu m - 2\mu + 2\mu p - 2\lambda \sum_0^\infty k \chi(k) p_{kk}, \tag{29}$$

$$\lambda p_{12} - \mu' p_{12} + \mu' p_{.3} + \lambda p_{.1} - \lambda p_{01} - \lambda w(1) p_{11} + \lambda p_{02} + \lambda p_{12} + \lambda w(2) p_{22} = 0, \tag{31}$$

$$D_{s's'}(1,1) = \lambda w' + 2m'\mu' - 2\mu' + 2\mu' p' + 2\lambda \sum_0^\infty k \chi(k) p_{kk}, \tag{33}$$

$$-(\lambda - \mu') p_{.1} m_1 + \mu' p_{.2} m_2 - \mu p_{.1} + \mu p_{01} + \lambda(p_{01} + 2w(1) p_{11} + p' m_0) = 0, \tag{35}$$

$$-(\lambda - \mu) p_{1.} m_1' + \mu p_{2.} m_2' - \mu' p_{1.} + \mu' p_{10} + \lambda(p_{10} + 2\chi(1) p_{11} + p m_0') = 0. \tag{36}$$

Equations (15), (26), (27), (30), (32) and (34) represent relations that are invariant with respect to $w(x)$. The expression for $D_{ss'}(1,1)$ is also independent of $w(x)$.

From the form of the general generating function equation, it is obvious that any relation between λ, μ, μ', and $p_{ss'}$ found by differentiating twice or more with respect to one of the pair s, s' and setting the other equal to zero is unchanged by altering $w(x)$.

The further modifications that must be made are left to the reader.

The author wishes to thank Dr Haight for help received first as his student and later as his colleague.

Reference

Haight, F. A. (1958). Two queues in parallel. *Biometrika*, **45**, 401–10.

The characteristic function of Hermitian quadratic forms in complex normal variables

By G. L. Turin

Hughes Research Laboratories, Culver City, California

Introduction

Wooding (1956) has given an expression for the multivariate distribution of several complex normal variables of a special type often encountered in the analysis of random noise. In this same connexion, a problem of great interest is the determination of the distribution of real-valued (hence, Hermitian) quadratic forms in these variables. Such a distribution arises, for example, in the analysis of the performance of a radar receiver with a square-law detector and a post-detection filter, for the output of such a receiver is representable as a weighted sum of the squared moduli of complex normal variables of this special type (cf. Kac & Siegert, 1947). Again, the distribution arises in the computation of the probability that an ideal binary hypothesis-testing communications receiver will err in its decision concerning which of two signals was transmitted through a channel comprising both a random transmission medium, such as the ionosphere, and additive thermal noise (Turin, 1958, 1959).

A first step in the determination of the distribution of the quadratic form, that of obtaining its characteristic function, is easily accomplished through the use of Wooding's result. The distribution itself may then be obtained by Fourier transformation of the characteristic function.

Although some of the results presented here have been derived elsewhere by use of a real-variable approach, it is felt that the complex-variable formulation, which arises so naturally in noise analysis, will be of interest in its own right, as well as in providing a simplified derivation of the characteristic function.

Derivation of the characteristic function

Let v_n $(n = 1, \ldots, N)$ be a set of complex random variables, the real and imaginary parts of which, x_n, y_n, are normally distributed and satisfy the relationships

$$E[(x_m - \bar{x}_m)(x_n - \bar{x}_n)] = E[(y_m - \bar{y}_m)(y_n - \bar{y}_n)] \tag{1}$$

and

$$E[(x_m - \bar{x}_m)(y_n - \bar{y}_n)] = -E[(x_n - \bar{x}_n)(y_m - \bar{y}_m)], \tag{2}$$

where $\bar{x}_n = E(x_n)$ and $\bar{y}_n = E(y_n)$. Denote by V the column matrix formed from the v_n, and let $\bar{V} = E(V)$ be the matrix of the (complex) means and $L = E[(V - \bar{V})(V - \bar{V})'^*]$ the (complex) covariance matrix, assumed non-singular. Further, let Q be any Hermitian matrix. Then it is found that the characteristic function of the (real) quadratic form

$$f = V'^* Q V \tag{3}$$

is given by

$$\phi(t) = |I - itLQ|^{-1} \exp\{-\bar{V}'^* L^{-1}[I - (I - itLQ)^{-1}]\,\bar{V}\}, \tag{4a}$$

where I is the unit matrix.

The quadratic form in the exponent of (4*a*) may be diagonalized by use of the transformation $\bar{V} = U_1 M^{-1} U_2 D$, where U_1 is the normalized modal matrix of L^{-1}, $M^2 = U_1'^* L^{-1} U_1$, and U_2 is the normalized modal matrix of $M^{-1} U_1'^* Q U_1 M^{-1}$. Then, alternatively,

$$\phi(t) = \prod_1^N \frac{\exp\left\{\dfrac{it\lambda_n |d_n|^2}{1 - it\lambda_n}\right\}}{1 - it\lambda_n}, \tag{4b}$$

where the d_n are the elements of D, and the λ_n are the eigenvalues of LQ (cf. Kac & Siegert, 1947).

To prove (4*a*), one first notes that the characteristic function of the random variable f is defined as

$$\phi(t) \equiv E(e^{itf}) = \int_V p(V) \exp(itV'^* Q V)\, dV, \tag{5}$$

where $p(V)$ is the multivariate density distribution of the N complex normal variables v_n, and

$$dV = \prod_1^N dx_n dy_n.$$

Wooding's expression for $p(V)$, modified to include explicitly the case for which $E(v_n) \neq 0$ (and thus including, in noise analysis, the case of a non-random or 'signal' component), is

$$p(V) = \pi^{-N} |L|^{-1} \exp[-(V - \bar{V})'^* L^{-1} (V - \bar{V})], \tag{6}$$

the exponent of which may be written

$$-\bar{V}'^* L^{-1} \bar{V} - V'^* L^{-1} V + \bar{V}'^* L^{-1} V + V'^* L^{-1} \bar{V}. \tag{7}$$

Since the covariance matrix L is clearly Hermitian, L^{-1} is Hermitian, and $(L^{-1})' = (L^{-1})^*$. Hence it follows that

$$\begin{aligned} \bar{V}'^* L^{-1} V &= [\bar{V}'(L^{-1})^* V^*]^* \\ &= [\bar{V}'(L^{-1})' V^*]^* \\ &= [V'^* (L^{-1})\, \bar{V}]^*, \end{aligned} \tag{8}$$

and (7) becomes

$$-\bar{V}'^* L^{-1} \bar{V} - V'^* L^{-1} V + 2\,\mathrm{Re}(\bar{V}'^* L^{-1} V). \tag{9}$$

Using (6) and (9), (5) becomes

$$\phi(t) = \pi^{-N} |L|^{-1} \exp(-\bar{V}'^* L^{-1} \bar{V}) \int_V \exp[2\,\mathrm{Re}(\bar{V}'^* L^{-1} V) - V'^* (L^{-1} - itQ)\, V]\, dV. \tag{10}$$

Now, it is easy to show by a generalization of a proof by Cramér (1951, pp. 118–20) that if A and V are any (N-element) column matrices and B and C are $(N \times N)$ Hermitian matrices,

$$\int_V \exp[\mathrm{Re}(A'^* V) - V'^* (B + iC)\, V]\, dV = \pi^N |B + iC|^{-1} \exp[\tfrac{1}{4} A'^* (B + iC)^{-1} A], \tag{11}$$

provided that B is positive definite. Since L is a non-singular covariance matrix, it is positive definite

(Cramér, 1951, p. 295); hence, L^{-1} is positive definite, and (11) may be applied to (10). The integral in (10) then becomes

$$\pi^N |L^{-1} - itQ|^{-1} \exp[\overline{V}'^* L^{-1}(L^{-1} - itQ)^{-1} L^{-1} \overline{V}], \tag{12}$$

which, inserted into (10), easily yields (4*a*).

Equation (4*b*) may alternatively be derived from an intermediate step in the proof of (11), which involves the simultaneous diagonalization of B and C (i.e. L^{-1} and Q). Equivalently, as a referee has noted, this diagonalization is to be regarded as reducing the form f of (3) to a linear combination of independent, non-central χ^2 variables of two degrees of freedom.

Evaluation of the density distribution

The density distribution of f,

$$p(f) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \phi(t)\, e^{-ift} dt, \tag{13}$$

is generally difficult to obtain precisely, because of the complicated nature of the exponential factor in $\phi(t)$. However, in the special case in which the variables v_n have zero means (corresponding, in noise problems, to the absence of a signal component), $\overline{V} = 0$, and (4*a*) and (4*b*) reduce to

$$\phi(t) = |I - itLQ|^{-1} = \prod_1^N (1 - it\lambda_n)^{-1}. \tag{14}$$

This is the reciprocal of an Nth degree polynomial in (it), the zeros of which lie at the reciprocals of the eigenvalues of the matrix LQ. The singularities of $\phi(t)$ in this case hence consist solely in a finite number of finite-order poles; once the eigenvalues of LQ are found, $p(f)$ is therefore readily determined through the use of the residue theorem.

It is interesting to note that this appearance of poles in $\phi(t)$ when $\overline{V} = 0$ occurs because of the special constraints placed on the $2N$ real random variables x_n, y_n by (1) and (2). It has been shown by Turin (1956) that if one considers more generally a quadratic form of unconstrained real variables Z with covariance matrix M, one obtains the characteristic function

$$\phi(t) = |I - 2itMR|^{-\frac{1}{2}} \exp\{-\tfrac{1}{2}\overline{Z}' M^{-1}[I - (I - 2itMR)^{-1}]\overline{Z}\}, \tag{15}$$

where R is the (real) matrix of the quadratic form; this expression may of course be put into a diagonal form similar to that of (4*b*). Clearly, in the special case when the real variables Z are constrained in the fashion prescribed by (1) and (2), the eigenvalues of MR must be of even multiplicity so that the necessary equality of (15) and (4) will obtain.† In general, however, such a pairing of eigenvalues does not occur. Hence, when $\overline{Z} = 0$ in (15) (cf. Whittle, 1951), $\phi(t)$ typically displays branch points among its singularities, rather than poles only, and the evaluation of $p(f)$ from $\phi(t)$ becomes correspondingly more difficult (see Grenander, Pollak & Slepian, 1958).

References

Cramér, H. (1951). *Mathematical Methods of Statistics*. Princeton University Press.

Grenander, U., Pollak, H. O. & Slepian, D. (1958). The distribution of quadratic forms in normal variates: a small sample theory with applications to spectral analysis. Unpublished paper.

Hoffman, W. C. (1953). A statistic associated with the joint distribution of n successive amplitudes. Ph.D. Thesis, Department of Mathematics, University of California at Los Angeles.

Kac, M. & Siegert, A. J. F. (1947). On the theory of noise in radio receivers with square-law detectors. *J. Appl. Phys.*, **18**, 383–97.

Turin, G. L. (1956). Communication through noisy, random multi-path channels. Lincoln Laboratory, Massachusetts Institute of Technology. Technical Report 116. See also, 1956 National Convention Record, Institute of Radio Engineers, part 4, pp. 154–66.

Turin, G. L. (1958). Error probabilities for binary symmetric ideal reception through nonselective slow fading and noise. *Proc. Inst. Radio Engrs, N.Y.*, **46**, 1603–19.

Turin, G. L. (1959). Some computations of error rates for selectively fading multipath channels. Proceedings of the 1959 National Electronics Conference, Chicago, Illinois, 12–14 October 1959. To be published.

Whittle, P. (1951). *Hypothesis Testing in Time Series Analysis*. Uppsala: Almquist and Wiksells, AB.

Wooding, R. A. (1956). The multivariate distribution of complex normal variables. *Biometrika*, **43**, 212–5.

† See Hoffman (1953) for a direct proof of the even multiplicity of the eigenvalues of MR in the constrained case.

Some numerical results for waiting times in the queue $E_k/M/1$

BY C. BURROWS
University College London

Kendall (1953) gave several equilibrium results for the queue represented by $GI/M/s$ (i.e. general independent input, exponential service-time distribution, s ($\geqslant 1$) servers) with the simple 'first come, first served' queue discipline. It is found that the equilibrium distribution of queue length is geometric except for modifications to the first $(s-1)$ terms, the common ratio for the geometric part of the series being λ which satisfies the equation

$$\lambda = \int_0^\infty e^{-(1-\lambda)u/\rho}\, dA(u) \quad (0 < \lambda < 1). \tag{1}$$

In this expression $A(u)$ is the cumulative distribution of the inter-arrival times adjusted so as to have a mean of unity, and ρ is the traffic density, i.e. the ratio of the mean service time divided by s, to the mean inter-arrival time.

In this note some numerical results are tabulated for the situation where the inter-arrival distribution is of the χ^2 form with ν degrees of freedom, so that

$$dA(u) = (\tfrac{1}{2}\nu)^{\frac{1}{2}\nu} \frac{1}{\Gamma(\frac{1}{2}\nu)} u^{\frac{1}{2}\nu-1} e^{-\frac{1}{2}u\nu} du.$$

If we take $\nu = 2k$ then, when $k = 1$, $dA(u)$ is an exponential distribution and as k tends to ∞, $dA(u)$ approaches a deterministic form. Thus in Kendall's notation the system is of the form $E_k/M/s$.

Equation (1) gives

$$\lambda = \int_0^\infty e^{-(1-\lambda)u/\rho}\, k^k \frac{1}{\Gamma(k)} u^k e^{-uk} du$$

$$= \left[\frac{\rho k}{\rho^{k+1} - \lambda}\right]^k.$$

This equation may be solved by iteration and values of λ are given in Table 1 for values of k and ρ.

Table 1. *Values of* λ, *the queue parameter*

	ρ									
$k-1$	0·05	0·1	0·2	0·3	0·4	0·5	0·6	0·7	0·8	0·9
0	0·0500	0·1000	0·2000	0·3000	0·4000	0·5000	0·6000	0·7000	0·8000	0·9000
0·5	·0122	·0504	·1298	·2218	·3215	·4267	·5359	·6484	·7635	·8808
1	$\cdot 0^3839$	·0292	·0938	·1780	·2753	·3820	·4958	·6155	·7399	·8682
2	$\cdot 0^3223$	·0127	·0590	·1319	·2239	·3305	·4485	·5759	·7109	·8526
3	$\cdot 0^3774$	$\cdot 0^3680$	·0430	·1083	·1963	·3019	·4216	·5529	·6939	·8433
4	$0 \cdot 0^3321$	$0 \cdot 0^3417$	0·0341	0·0942	0·1792	0·2838	0·4043	0·5379	0·6828	0·8372
5	$\cdot 0^3151$	$\cdot 0^3281$	·0285	·0849	·1675	·2712	·3922	·5274	·6749	·8328
10	$\cdot 0^4112$	$\cdot 0^3818$	·0172	·0642	·1405	·2414	·3629	·5017	·6553	·8229
15	$\cdot 0^5232$	$\cdot 0^3424$	·0136	·0567	·1302	·2292	·3512	·4913	·6474	·8175
20	$\cdot 0^5791$	$\cdot 0^3281$	·0118	·0528	·1247	·2235	·3450	·4857	·6432	·8151
∞	$0 \cdot 0^7206$	$0 \cdot 0^4454$	$0 \cdot 0^2697$	0·0409	0·1074	0·2032	0·3242	0·4670	0·6286	0·8069

The values in the row $k-1 = \infty$ correspond to Table 2 in Kendall (1953, p. 351). The table shows that as k increases and the inter-arrival time distribution departs more and more from the exponential, λ decreases so that the queue length distribution has decreased probabilities for its tail values.

The distribution of waiting time (w) (i.e. the time spent in the queue up to the start of service) is exponential except for a probability concentration at $w = 0$, and for the case where we have only one server ($s = 1$) Kendall (1953) gives for the expected waiting time

$$R = \frac{\mathscr{E}(w)}{\mathscr{E}(v)} = \frac{\lambda}{1-\lambda}.$$

Here R is the ratio of expected time spent waiting for service to the expectation of the time spent in service (v). This may be regarded as a figure of demerit for the system. When the inter-arrival time distribution is of the exponential form we have

$$R_M = \frac{\rho}{1-\rho},$$

and in Table 2 values of R/R_M are tabulated.

Table 2. *Values of* R/R_M

	ρ										
$k-1$	0·05	0·1	0·2	0·3	0·4	0·5	0·6	0·7	0·8	0·9	R'/R_M
0·5	0·2347	0·4775	0·5968	0·6650	0·7107	0·7443	0·7698	0·7903	0·8071	0·8210	0·8333
1	·1608	·2705	·4140	·5052	·5699	·6181	·6555	·6861	·7112	·7319	·7500
2	·0425	·1154	·2509	·3544	·4327	·4937	·5421	·5820	·6148	·6427	·6667
3	·0147	·0615	·1796	·2835	·3663	·4325	·4859	·5300	·5667	·5978	·6250
4	$0{\cdot}0^3609$	0·0377	0·1410	0·2427	0·3275	0·3963	0·4525	0·4989	0·5382	0·5714	0·6000
5	$\cdot0^3287$	·0254	·1172	·2165	·3018	·3721	·4302	·4783	·5190	·5534	·5833
10	$\cdot0^3214$	$\cdot0^2737$	·0700	·1600	·2452	·3182	·3797	·4315	·4753	·5163	·5455
15	$\cdot0^4440$	$\cdot0^3382$	·0552	·1401	·2246	·2974	·3609	·4139	·4590	·4977	·5313
20	$\cdot0^4150$	$\cdot0^3253$	·0479	·1300	·2137	·2878	·3511	·4047	·4507	·4898	·5238
∞	$0{\cdot}0^6392$	$0{\cdot}0^3409$	0·0281	0·0995	0·1805	0·2550	0·3198	0·3755	0·4231	0·4643	0·5000
R_M	0·0526	0·1111	0·2500	0·4286	0·6667	1·0000	1·5000	2·3333	4·0000	9·0000	

We may compare these results with those obtained when the inter-arrival time distribution is exponential, while the service-time distribution is not necessarily so. From Kendall (1951) we have

$$R' = \frac{\mathscr{E}(w)}{\mathscr{E}(v)} = \frac{\rho}{2(1-\rho)}\left[1+\mathrm{var}\left(\frac{v}{b}\right)\right],$$

where var (v) is the variance of the service-time distribution which has mean b. R' is a minimum when var (v) is zero, i.e. all service times are of equal length, and then R' is half of the value R_M. The ratio R'/R_M is independent of ρ and in the extreme right-hand column of Table 2 values of the ratio are given for the case where the service-time distribution is of the χ^2 form with $2k$ degrees of freedom.

By comparing the body of Table 2 giving the ratio R/R_M with the right-hand column giving R'/R_M we see that a change in arrival distribution can bring about a larger saving in waiting time than a change in service-time distribution. This is especially true when the traffic density is small.

References

Kendall, D. G. (1951). Some problems in the theory of queues. *J. R. Statist. Soc.* B, **13**, 151–85.

Kendall, D. G. (1953). Stochastic processes occurring in the theory of queues and their analysis by the method of the imbedded Markov chain. *Ann. Math. Statist.* **24**, 338–54.

A formula for the curvature of the likelihood surface of a sample drawn from a distribution admitting sufficient statistics

By B. RAJA RAO

Department of Mathematics and Statistics, University of Poona, India

1. Summary

In this paper is given a formula for the curvature of the 'likelihood surface' (see § 3) of a sample drawn from a distribution admitting sufficient statistics for the parameters, at the point represented by the maximum likelihood estimates (m.l.e.) of the parameters. More precisely it is shown in the case of two parameters that the negative of the trace and the determinant of Fisher's information matrix, respectively, measure the first and second (Gaussian) curvatures of the 'likelihood surface' at the point represented by the m.l.e.'s of the parameters. In the case of m parameters an expression for the Riemann

curvature of the 'likelihood hypersurface' (see § 4) of a sample at the point represented by the m.l.e.'s is obtained in terms of the elements of Fisher's information matrix. An expression for Riemann's curvature invariant is also obtained. A large sample interpretation is given.

2. Introduction

As Durbin & Kendall (1951) remark, apart from the use of a multi-dimensional space by Neyman and Pearson in connexion with fitting theoretical distributions, by Hotelling and Fisher in a discussion of maximum likelihood estimators, and apart from a remark by Huzurbazar (1949) concerning the radius of curvature of a likelihood curve, there seem to have been few attempts to consider estimation as a geometrical problem. In the following paper we shall make use of differential and Riemannian geometries to obtain the curvatures of a 'likelihood surface' and a 'likelihood hypersurface'.

Let $f(x, \theta_j)$ be the probability density function (p.d.f.) of a distribution involving a set of m unknown parameters. Under certain general regularity conditions, Koopman (1936) has shown that the necessary and sufficient condition that $f(x, \theta_j)$ should admit a set of m jointly sufficient statistics is that $f(x, \theta_j)$ be in the form

$$f(x, \theta_j) = \exp\left\{\sum_{\nu=1}^{m} u_\nu(\theta_j)\, v_\nu(x) + A(x) + B(\theta_j)\right\}. \tag{1}$$

If $x_1, x_2, \ldots, x_n$ is a random sample of n independent observations from this distribution, we shall for convenience describe $L = \sum_{i=1}^{n} \log f(x_i, \theta_j)$ as the log-likelihood function of the sample. Let $\hat{\theta}_j, j = 1, 2, \ldots, m$ be a solution of the system of likelihood equations

$$\frac{\partial L}{\partial \theta_\nu} = 0 \quad (\nu = 1, 2, \ldots, m),$$

so that the $\hat{\theta}_j$ are the m.l.e.'s of the θ_j. Then Huzurbazar (1949) has shown that the above system has a unique consistent solution $\theta_j = \hat{\theta}_j, j = 1, 2, \ldots, m$ for all samples of any size, that the solution $\theta_j = \hat{\theta}_j$ makes the likelihood an absolute maximum and has obtained the interesting property

$$\left[\left(-\frac{\partial^2 L}{\partial \theta_r\, \partial \theta_s}\right)_{\theta_j = \hat{\theta}_j}\right] = \left[\left\{E\left(-\frac{\partial^2 L}{\partial \theta_r\, \partial \theta_s}\right)\right\}_{\theta_j = \hat{\theta}_j}\right] = [\hat{I}_{rs}], \tag{2}$$

where $[I_{rs}]$ is Fisher's information matrix and the circumflex ($\hat{}$) indicates that the value I_{rs} is evaluated at $\theta_j = \hat{\theta}_j$.

In this paper we shall obtain a formula for the first and second curvatures of the 'likelihood surface' in the case of two parameters and a formula for the Riemann curvature of the 'likelihood hypersurface' represented by the likelihood function of the sample at the point given by the m.l.e.'s of the θ_j.

3. Case of two parameters

For the sake of simplicity we shall first consider the case of two parameters. Let $x_1, x_2, \ldots, x_n$ be a random sample of size n from a distribution of Koopman's form (1) depending on the two parameters θ_1, θ_2. If we consider a fixed sample and variation in the parameter space, the locus described by $L = L(\theta_1, \theta_2)$, in which one co-ordinate is written as a function of the others, will represent a surface in the parameter space, which we might call the 'likelihood surface' of the sample, whose equation is now said to be in Monge's form.

One interesting question which immediately arises is why should we consider samples drawn from only distributions admitting sufficient statistics? Why should not we sample from any distribution? First of all, we need regular conditions and our results will not hold when the range of the distribution depends on the parameters to be estimated. Secondly, the metric L for a distribution admitting sufficient statistics is itself an interesting one and it is quite well known that this form of the metric is almost inevitable in measuring the 'distance' between two populations. The role of sufficient statistics seems to depend on the fact that $\partial L/\partial\theta =$ function of θ and the sufficient statistics. Further, in our treatment we require Huzurbazar's result, given by equation (2), that at the m.l.e. point $\theta = \hat{\theta}$ (to take the instance of a single parameter) the value of $\partial^2 L/\partial\theta^2$ when $\theta = \hat{\theta}$ is the same as its expected value when $\theta = \hat{\theta}$, which is true only when there is a sufficient statistic for θ. Such a point $\theta = \hat{\theta}$ on the likelihood surface is known to be unique. In the case of a general distribution this is not possible.

We shall now obtain the first and the second (Gaussian) curvatures of the likelihood surface at the point $(\hat{\theta}_1, \hat{\theta}_2)$ represented by the m.l.e.'s of θ_1 and θ_2.

In the usual notation (Weatherburn, 1927, p. 75) we have

$$\left.\begin{aligned} p &= \frac{\partial L}{\partial \theta_1} \\ q &= \frac{\partial L}{\partial \theta_2} \end{aligned}\right\} = 0$$

at, and only at, $(\hat{\theta}_1, \hat{\theta}_2)$ as shown by Huzurbazar. Also at the point $(\hat{\theta}_1, \hat{\theta}_2)$ we have

$$E = 1, \quad F = 0, \quad G = 1, \quad H^2 = EG - F^2 = 1$$

(note that the parametric curves $\theta_1 =$ constant and $\theta_2 =$ constant are orthogonal at $(\hat{\theta}_1, \hat{\theta}_2)$),

$$e = \partial^2 L/\partial \theta_1^2, \quad f = \partial^2 L/\partial \theta_1 \partial \theta_2, \quad g = \partial^2 L/\partial \theta_2^2. \tag{3}$$

(We have used e, f, g for the second fundamental magnitudes instead of L, M, N in Weatherburn's notation to avoid confusion, for L is already used to denote the log-likelihood function of the sample.) Note that by formula (2)

$$\begin{aligned} f = \partial^2 L/\partial \theta_1 \partial \theta_2 &= \left\{ E\left(+ \frac{\partial^2 L}{\partial \theta_1 \partial \theta_2} \right) \right\}_{\theta_j = \hat{\theta}_j} \\ &= 0 \end{aligned}$$

if θ_1 and θ_2 are orthogonal (Jeffreys, 1948), so that in such a case the parametric curves at $(\hat{\theta}_1, \hat{\theta}_2)$ are also the lines of curvature. Thus at $(\hat{\theta}_1, \hat{\theta}_2)$

$$\begin{aligned} T^2 &= eg - f^2 \\ &= (\partial^2 L/\partial \theta_1^2)(\partial^2 L/\partial \theta_2^2) - (\partial^2 L/\partial \theta_1 \partial \theta_2)^2 \\ &= \begin{vmatrix} -\partial^2 L/\partial \theta_1^2 & -\partial^2 L/\partial \theta_1 \partial \theta_2 \\ -\partial^2 L/\partial \theta_1 \partial \theta_2 & -\partial^2 L/\partial \theta_2^2 \end{vmatrix} \\ &= \det \begin{bmatrix} \hat{I}_{11} & \hat{I}_{12} \\ \hat{I}_{12} & \hat{I}_{22} \end{bmatrix} \quad \text{(again by formula (2))} \\ &= \det [\hat{I}_{rs}], \end{aligned} \tag{4}$$

say, where $[I_{rs}]$ is Fisher's information matrix. Then we have at the point $(\hat{\theta}_1, \hat{\theta}_2)$:

$$\begin{aligned} \text{the first curvature} \quad J &= (Eg - 2Ff + Ge)/H^2 \\ &= \partial^2 L/\partial \theta_1^2 + \partial^2 L/\partial \theta_2^2 \\ &= -\{I_{11}(\hat{\theta}_1, \hat{\theta}_2) + I_{22}(\hat{\theta}_1, \hat{\theta}_2)\} \\ &= -\operatorname{trace} [\hat{I}_{rs}], \end{aligned}$$

and

$$\begin{aligned} \text{the second curvature} \quad K &= T^2/H^2 \\ &= \det [\hat{I}_{rs}], \end{aligned}$$

from equation (4). Thus the negative of the trace of the information matrix and the determinant of the information matrix respectively measure the first and the second curvatures of the likelihood surface represented by the likelihood function at $(\hat{\theta}_1, \hat{\theta}_2)$. Or in other words we note that at the point $(\hat{\theta}_1, \hat{\theta}_2)$ the curvatures $-J$ and K are respectively the sum and product of the latent roots of the information matrix.

For large samples we have another interpretation. The variance-covariance matrix of $\hat{\theta}_1$ and $\hat{\theta}_2$ is then $[I_{rs}(\theta_1, \theta_2)]^{-1}$, the estimate of the determinant of which is $\{\det [\hat{I}_{rs}]\}^{-1} = 1/K$. Thus the reciprocal of the Gaussian curvature of the likelihood surface at the point $(\hat{\theta}_1, \hat{\theta}_2)$ gives an estimate of the generalized variance of the m.l.e.'s in large samples.

4. Case of several parameters

Let us now consider the general case of m parameters $\theta_1, \theta_2, \ldots, \theta_m$. Let S_{m+1} be a Euclidean space of $m+1$ dimensions referred to the co-ordinates ϕ_α, $\alpha = 1, 2, \ldots, m+1$, and having the metric $a_{\alpha\beta} d\phi_\alpha d\phi_\beta$. Points of S_{m+1} whose co-ordinates are expressible as functions of m independent variables $\theta_i, i = 1, 2, \ldots, m$ are said to constitute a V_m immersed in S_{m+1}. The space V_m is then said to be a hypersurface. Let the metric for the V_m be denoted by $g_{ij} d\theta_i d\theta_j$. Then the coefficients of the two forms are connected by the relations

$$g_{ij} = a_{\alpha\beta} \frac{\partial \phi_\alpha}{\partial \theta_i} \frac{\partial \phi_\beta}{\partial \theta_j}. \tag{5}$$

Consider now the particular hypersurface defined by

$$\left.\begin{aligned} \phi_1 &= \theta_1, \\ \phi_2 &= \theta_2, \\ &\vdots \quad \vdots \\ \phi_m &= \theta_m, \\ \phi_{m+1} &= L(\theta_1, \theta_2, \ldots, \theta_m), \end{aligned}\right\} \tag{6}$$

where $L(\theta_1, \theta_2, \ldots, \theta_m)$ is the log-likelihood function of the sample. We might call it the 'likelihood hypersurface' of the sample. In such a case we achieve a simplification by choosing the co-ordinates ϕ_α in S_{m+1} as Euclidean co-ordinates, so that

$$a_{\alpha\alpha} = 1, \quad a_{\alpha\beta} = 0, \quad \alpha \neq \beta;$$

then the coefficients g_{ij} of the fundamental form for V_m are given by

$$\begin{aligned} g_{ij} = \frac{\partial \phi_\alpha}{\partial \theta_i}\frac{\partial \phi_\alpha}{\partial \theta_j} &= \sum_{\alpha=1}^{m} \frac{\partial \phi_\alpha}{\partial \theta_i}\frac{\partial \phi_\alpha}{\partial \theta_j} + \frac{\partial \phi_{m+1}}{\partial \theta_i}\frac{\partial \phi_{m+1}}{\partial \theta_j} \\ &= \sum_{\alpha=1}^{m} \delta_i^\alpha \delta_j^\alpha + \frac{\partial L}{\partial \theta_i}\frac{\partial L}{\partial \theta_j}. \end{aligned}$$

Therefore we have the metric

$$\left.\begin{aligned} g_{ij} &= \frac{\partial L}{\partial \theta_i}\frac{\partial L}{\partial \theta_j} \quad (i \neq j), \\ g_{ii} &= 1 + \left(\frac{\partial L}{\partial \theta_i}\right)^2. \end{aligned}\right\} \tag{7}$$

Let now $(\hat{\theta}_1, \hat{\theta}_2, \ldots, \hat{\theta}_m)$ be a solution of the system of likelihood equations

$$\frac{\partial L}{\partial \theta_\nu} = 0 \quad (\nu = 1, 2, \ldots, m). \tag{8}$$

Then at the point $(\hat{\theta}_1, \hat{\theta}_2, \ldots, \hat{\theta}_m)$ the metric g_{ij} becomes, say $\hat{g}_{ij}$, where

$$\left.\begin{aligned} \hat{g}_{ij} &= 0 \quad (i \neq j), \\ \hat{g}_{ii} &= 1, \end{aligned}\right\} \tag{9}$$

so that the reciprocal tensor $\hat{g}^{ij}$ is

$$\left.\begin{aligned} \hat{g}^{ij} &= 0 \quad (i \neq j), \\ \hat{g}^{ii} &= 1. \end{aligned}\right\} \tag{10}$$

We shall now obtain an expression for the Riemann curvature of the likelihood hypersurface V_m at the point $(\hat{\theta}_1, \hat{\theta}_2, \ldots, \hat{\theta}_m)$ for the orientation determined by two unit vectors A^i and B^j. This is defined by (see Weatherburn, 1950, p. 117)

$$K = \frac{R_{rjnp} A^r A^n B^j B^p}{(g_{rn}g_{jp} - g_{rp}g_{jn}) A^r A^n B^j B^p}, \tag{11}$$

where the quantities R_{rjnp} are the components of Riemann's curvature tensor and are defined by

$$R_{rjnp} = \frac{\partial}{\partial \theta_n}\left(g_{rl}\begin{Bmatrix} l \\ jp \end{Bmatrix}\right) - \begin{Bmatrix} l \\ jp \end{Bmatrix}\frac{\partial g_{rl}}{\partial \theta_n} - \frac{\partial}{\partial \theta_p}\left(g_{rl}\begin{Bmatrix} l \\ jn \end{Bmatrix}\right) + \begin{Bmatrix} l \\ jn \end{Bmatrix}\frac{\partial g_{rl}}{\partial \theta_p} + g_{rl}\begin{Bmatrix} l \\ ns \end{Bmatrix}\begin{Bmatrix} s \\ jp \end{Bmatrix} - g_{rl}\begin{Bmatrix} l \\ ps \end{Bmatrix}\begin{Bmatrix} l \\ jn \end{Bmatrix},$$

where $[k, ij]$ (see below) and $\begin{Bmatrix} k \\ ij \end{Bmatrix}$ are Christoffel symbols of the first and second kinds respectively defined by

$$[k, ij] = \frac{1}{2}\left(\frac{\partial g_{jk}}{\partial x^i} + \frac{\partial g_{ik}}{\partial x^j} - \frac{\partial g_{ij}}{\partial x^k}\right)$$

and

$$\begin{Bmatrix} k \\ ij \end{Bmatrix} = g^{kh}[h, ij].$$

Thus we may write

$$R_{rjnp} = T_1 + T_2, \tag{12}$$

where

$$T_1 = \frac{1}{2}\left(\frac{\partial^2 g_{rp}}{\partial \theta_j \partial \theta_n} + \frac{\partial^2 g_{jn}}{\partial \theta_r \partial \theta_p} - \frac{\partial^2 g_{rn}}{\partial \theta_j \partial \theta_p} - \frac{\partial^2 g_{jp}}{\partial \theta_r \partial \theta_n}\right) \tag{13}$$

and

$$T_2 = g^{ts}([jn, s][rp, t] - [jp, s][rn, t]). \tag{14}$$

Using the metric defined by the equations (7) we obtain

$$T_1 = \frac{\partial^2 L}{\partial\theta_r \partial\theta_n}\frac{\partial^2 L}{\partial\theta_j \partial\theta_p} - \frac{\partial^2 L}{\partial\theta_r \partial\theta_p}\frac{\partial^2 L}{\partial\theta_j \partial\theta_n} \tag{15}$$

and

$$T_2 = -g^{ts}\frac{\partial L}{\partial\theta_t}\frac{\partial L}{\partial\theta_s}T_1. \tag{16}$$

Evaluating the various quantities at the point $(\theta_1, \theta_2, \ldots, \theta_m)$ we find, since $(T_2)_{\theta_j=\hat{\theta}_j} = 0$,

$$\begin{aligned}\hat{R}_{rjnp} &= \frac{\partial^2 L}{\partial\theta_r \partial\theta_n}\frac{\partial^2 L}{\partial\theta_j \partial\theta_p} - \frac{\partial^2 L}{\partial\theta_r \partial\theta_p}\frac{\partial^2 L}{\partial\theta_j \partial\theta_n} \\ &= (-\hat{I}_{rn})(-\hat{I}_{jp}) - (-\hat{I}_{rp})(-\hat{I}_{jn}) \\ &= \begin{vmatrix} \hat{I}_{rn} & \hat{I}_{rp} \\ \hat{I}_{jn} & \hat{I}_{jp} \end{vmatrix},\end{aligned}$$

by formula (2). Using the formula (11) and choosing the vectors A^i and B^j as orthogonal (so that the denominator of K is unity) we find that the Riemann curvature of the likelihood hypersurface V_m at the point $(\theta_1, \theta_2, \ldots, \theta_m)$ associated with the orthogonal unit vectors A^i and B^j is given by

$$K = \sum_{r,j,n,p} \begin{vmatrix} \hat{I}_{rn} & \hat{I}_{rp} \\ \hat{I}_{jn} & \hat{I}_{jp} \end{vmatrix} A^r A^n B^j B^p. \tag{17}$$

We are now faced by the problem: how many (distinct) non-zero terms are there in the above quadruple series? It is easy to see that it is the same as the number of distinct arithmetical non-vanishing components of the tensor R_{rjnp}, and it can be shown that the required number is $\frac{1}{12}m^2(m^2-1)$.

Since the Riemann curvature K determined by (17) is associated with the particular orientation defined by the vectors A^i and B^j, it is convenient to consider Riemann's curvature invariant R, which is independent of the particular orientation. It is defined by

$$R = g^{pr}g^{jn}R_{rjnp},$$

so that at the point $(\theta_1, \theta_2, \ldots, \theta_m)$ it gives

$$\begin{aligned}R &= \hat{g}^{pr}\hat{g}^{jn}\hat{R}_{rjnp} \\ &= \hat{R}_{rnnr}, \quad \text{from (10)} \\ &= \begin{vmatrix} \hat{I}_{rn} & \hat{I}_{rr} \\ \hat{I}_{nn} & \hat{I}_{rn} \end{vmatrix} \\ &= -\sum_{r,n=1}^{m} \begin{vmatrix} \hat{I}_{rr} & \hat{I}_{rn} \\ \hat{I}_{rn} & \hat{I}_{nn} \end{vmatrix}.\end{aligned}$$

Thus the sum of the determinants of all the possible $\frac{1}{2}m(m-1)$ leading 2×2 submatrices of the information matrix measures the negative of Riemann curvature invariant R at the point $(\hat{\theta}_1, \hat{\theta}_2, \ldots, \hat{\theta}_m)$ represented by the m.l.e.'s of $\theta_1, \theta_2, \ldots, \theta_m$. As in the case of two parameters we note that the Riemann curvature invariant at the point $(\hat{\theta}_1, \hat{\theta}_2, \ldots, \hat{\theta}_m)$ is the negative of the sum of the products of the latent roots of Fisher's information matrix taken two at a time.

As in § 3 we could give a large sample interpretation of our result.

I wish to record my great indebtedness to Prof. V. S. Huzurbazar for his constant encouragement during the course of this investigation, and to Dr (Mrs) M. Apte for some helpful discussions I have had, which have benefited me greatly in § 4. My grateful thanks are due to Prof. M. G. Kendall for his valuable suggestions and comments on an earlier draft of this paper which have added to its value.

References

Durbin, J. & Kendall, M. G. (1951). *Biometrika*, **38**, 150.
Huzurbazar, V. S. (1949). *Biometrika*, **36**, 71.
Jeffreys, H. (1948). *Theory of Probability* (2nd ed.). Oxford: Clarendon Press.
Koopman, B. O. (1936). *Trans. Amer. Math. Soc.* **39**, 399.
Weatherburn, C. E. (1927). *Differential Geometry*. Cambridge University Press.
Weatherburn, C. E. (1950). *Riemannian Geometry*. Cambridge University Press.

Reviews

Statistics of Extremes. By E. J. GUMBEL. New York: Columbia University Press. 1958. Pp. xx + 373. $15.00.

A saying, to be found equivalently in many languages besides English, states that 'a chain is as strong as its weakest link'. This aphorism states succinctly and accurately why statisticians, and especially perhaps statisticians advising engineers, are concerned with extreme value problems. E. J. Gumbel has been concerned for many years with the study of these and allied problems and has, in this present book, put down the fruits of his researches. The book, written by a specialist, is therefore one which would *a priori* command attention, and the contents discover to us that this attention is very much worth while.

Prof. Gumbel requires of his readers that they should have as prerequisites some calculus, some analysis and a little more than a smattering of mathematical statistics and probability. As delineated by the author the mathematical structure of the book is not difficult but the uninitiated will not be able to plunge into it in a carefree way. The book begins with a longish chapter (38 pages) in which the author sets out the tools he is going to use and seeks to orientate his readers so that they understand the blue-print from which he is going to work. Chapter 2 is on order statistics and Chapter 3 on the exact distribution of extremes. As is well known the study of exact distributions in any form of 'order' problem is, on the whole, an unprofitable proceeding so we turn in Chapter 4 to the analytic study of extremes, with asymptotic results following in Chapters 5, 6 and 7. The range has Chapter 8 to itself. There are 44 tables and over 100 graphs.

The emphasis, as far as the author's examples go, is on mechanical engineering problems (such as fatigue in metals) and on climatic difficulties which are really civil engineering problems (such as floods and droughts). But it would be a mistake to conclude that only engineers with a statistical slant could profitably read and learn. Anyone who is interested in extreme values will have to read this treatise in order to find out what has been done. In the limited field of mathematical statistics which this book delineates it is encyclopaedic and can be unhesitatingly and unreservedly recommended.

F. N. DAVID

International Journal of Abstracts: Statistical Theory and Method. Vol. 1, No. 1, ed. W. R. BUCKLAND. London and Edinburgh: Oliver and Boyd for the International Statistical Institute. 1959. Subscription £5 or $16 per annum

This new journal, to be issued quarterly, is a compilation of straight abstracts, of about 300 words, without comment. There are 160 abstracts in this number, printed two to a page (on one side only, so that they can be cut out and stuck on index cards if desired). They vary in length and precision depending to some extent on the length and originality of the papers abstracted and also on the degree to which the divers abstractors have digested what they have read. Since the whole relevant literature (nearly 400 journals) is scanned the price of the journal is reasonable and it provides, in abstract form at least, translations of Russian and Japanese papers for instance which are otherwise effectively unobtainable. Thus, for the first time, it will be possible for most statisticians to feel reasonably sure of being aware, at least in outline, of all recent work.

There is no table of contents nor index (but a running index is to be initiated at the end of the volume); here, the abstracts are classified into eleven groups of nine subsections, the papers being arranged alphabetically, according to author, in each group. Since many of the abstracts are highly condensed the effect of misprints is much more damaging than in other texts and it would have been worthwhile to take more trouble with the proof reading even at the price of delay and extra cost. Taken all in all, however, it is a product of remarkably high quality and we may count ourselves very well served if we continue to see its like.

D. E. BARTON

Testing Statistical Hypothesis. By E. L. LEHMANN. New York: John Wiley and Sons Inc.; London: Chapman and Hall. 1959. Pp. 369. 88*s*.

An outline of this book appeared in 1949, and in the years following parts of it were presented in lectures by the author. The whole field of testing statistical hypotheses is not covered in this volume; the main concern being the two-decision problem as it presents itself after the experiment has been set up. Other areas not fully explored are essentially the chi-square and likelihood ratio tests.

The first two chapters are introductory, the first introducing in detail the general decision problem, based on loss and risk considerations; in the second the background of probability as a part of the theory of measure and integration is outlined. This general decision approach is as emphasized, only to serve as basis, for throughout the remainder of the book the Neyman–Pearson formulation with the simpler notion of error instead of loss is used.

The next four chapters may be considered as representative of the book. With the mathematical framework of maximizing an integral subject to certain side conditions or, more specifically, based on the Neyman–Pearson fundamental lemmas, an extensive discussion of the theory of uniformly most powerful tests is given. Likewise the concept of uniformly most powerful tests is discussed in connexion with the class of unbiased tests, and invariance tests. This developed theory is then applied to various statistical problems, special attention being paid to the important exponential family and the closely related problem of confidence bounds; it is also shown that various existing tests have optimum properties within certain classes.

In the remaining two chapters the theory and underlying ideas of analysis of variance, chi-square tests, likelihood ratio tests and minimax procedures are briefly sketched.

Throughout use is made of numerous examples to illustrate and emphasize, and each chapter is concluded with an extensive set of classified problems. Worthy of mention also are the references which are printed in such a way that they give a good guide to the related literature. As to the presentation of the material, a slight lack of explanation may be felt in some instances. However, although the whole field of testing statistical hypotheses is not covered, the book makes clear the author's comprehensive study and research in this theory, and both the teacher and student in the field will find the book a most useful one.

C. F. CROUSE

Die axiomatischen Grundlagen einer allgemeinen Theorie des Messens. (Axiomatic foundations of the general theory of measurement). By J. PFANZAGL. Würzburg: Physica Verlag. 1959. Pp. 63. DM 14.

This is the first of a series of monographs from the Statistical Institute of the University of Vienna. It is especially appropriate to a pioneer publication in that it studies a problem which does not appear frequently in statistical literature. This is the problem of defining what are the essential properties of a 'measurement'. The author says that it is common practice to require that measurement puts every member of a set M into the 'strongest possible' relationship with a set of real numbers (the measurements): and that, in most cases, it is required that some form of 'additivity' condition should be satisfied. The author claims that this last condition is unnecessarily restrictive, and suggests replacing it by a general 'so-called metrical operation'—continuous, monotonic and one-to-one.

It is salutary to pay close attention, from time to time, to the true nature of measurements (the 'observational data' or 'sample values' of statistics). Such works as the present monograph remind us that the very concept of 'measurement' is capable of modification, and perform a useful service in drawing attention to fundamentals. Many statisticians will find the present work rather abstract, but an understanding of the author's aims and methods should repay the necessary effort.

N. L. JOHNSON

Statistical Estimates and Transformed Beta-Variables. By GUNNAR BLOM. New York: John Wiley and Sons, Inc.; Stockholm: Almquist and Wiksell; London: Chapman and Hall. 1958. Pp. 176. 40*s*.

This book is presumably a published form of a doctoral thesis. The author states that he began by being interested in estimation by means of order statistics and was led from this (or perhaps one should say through this) to an investigation of the theory of estimation. The first two chapters consist of

a mathematical discussion of the Fisherian theory of estimation with some generalization in the direction which the author is going to require for the rest of his researches. In Chapter 3 he moves away from the usual terminology of order statistic to that of transformed beta-variable. It is well known that any order statistic has a beta-distribution when expressed in terms of its cumulative distribution function. It is doubtful whether this new nomenclature adds clarity to our ideas on the subject but the author thinks that he has generalized the concept by this means.

From the change in nomenclature we go on to properties and approximate properties of transformed beta-variables, and of linear combinations of them. We then have the applications of the theory of linear estimation to these variables. This section of about 70 pages is important, particularly in connexion with the treatment of estimation in censored samples, a topic which has recently been receiving a certain amount of attention.

The book throughout is written in highly condensed fashion with due attention to mathematical rigour. The reader envisaged can only be a mathematical statistician since no clear picture emerges as to possible practical applications.

F. N. DAVID

Statistics: An Introduction. By D. A. S. FRASER. New York: John Wiley and Sons Inc.; London: Chapman and Hall. 1959. Pp. ix + 396. 54*s*.

This is yet another book on the mathematical theory of statistics. It is intended, the author tells us, as a text for a two-semester course for students with the prerequisite of a full year of calculus. (It might also have been added that some knowledge of matrix algebra is desirable.) The book follows the orthodox theme. The topics covered, in order of presentation, are discrete and continuous probability distributions, expectation techniques and generating functions, central limit theorem, single and two-way classification, χ^2 and F, estimation, testing hypotheses, confidence intervals, regression analysis, factorial designs with one, two and three factors, randomization, analysis of covariance, and fractional replication, sequential analysis and non-parametric methods. The degree of sophistication required is a little more than that for A. M. Mood, *Elementary Theory of Statistics*.

The level of exposition of this book is adequate and it will be found useful by mathematics students wanting to learn something of the mysteries of the statistical craft. If it does not add anything in the way of fresh theory or example to what has already been published this is not the fault of the author, but merely reflects the fact that a great many 'introductions' and 'elements' have already appeared. For those students who seek to buy a text-book on statistics for the first time this is a suitable book to buy. For those students who have already bought an introductory text-book stronger meat than this will probably be required.

F. N. DAVID

Operations Research. An Annotated Bibliography. Second edition. By JAMES H. BATCHELOR. Saint Louis, Missouri: Saint Louis University Press. 1959. $10.

The price of this book is likely to put it out of the reach of private individuals, but it can be said without hesitation that this work should be in any library concerned with its subject. It lists 4195 titles of papers or books concerned with operational research and in most cases gives an abstract of from 50 to 150 words (occasionally more) of the papers listed. The author refers in his introduction to paper 1321, by Merrill M. Flood in which it is said that 'operations research represents nothing more than a name for the post-war increase in the quantity and quality of scientific research on industrial management and operating problems', and in the light of this, completeness is impossible of attainment in such a bibliography. A notable omission is the absence of reference to the book by McGuire, Bechmann and Winsten on problems of transport, which is in many ways a model of how operational research should be carried out. And doubtless a further search would reveal further omissions. On the other hand, a notable inclusion is: 0060, E. D. Adrian, *The Physical Background of Perception*.

The value of the bibliography is greatly enhanced by an index, mainly by subject, which occupies 84 pages.

Mr Batchelor has put us all in his debt by his painstaking work.

G. A. BARNARD

Quality Control and Industrial Statistics. (Revised edition.) By ACHESON J. DUNCAN. Homewood, Illinois: Richard D. Irwin Inc. 1959. Pp. xxxiii + 946. $10.00.

This second enlarged and revised edition is to be very warmly welcomed. The main changes are the addition of 250 pages of text, the creation of a separate section for rectifying inspection, and an improved lay-out (with additions) for acceptance sampling by variables. For the rest, apart from some minor pruning and addition, and the compilation of more up-to-date reference lists, the text follows that of the first edition reviewed by Cox in *Biometrika*, volume 40. It is convenient to discuss the extra material (with the approximate number of extra pages given in brackets) under the headings given in the book.

Part II. Lot acceptance sampling plans. The enlarged section on sampling inspection by variables includes a description of the Mil. Std. 414, a good exposition of sequential probability ratio sampling plans, and a short chapter comparing sampling plans by the average sample numbers, and administrative features (40 pages). There is also a new chapter (9 pages) on the Barnard–Enters–Hamaker multiple sampling procedure.

Part III. Rectifying inspection. This part contains material which was intermingled with other sampling plans in the first edition, and the creation of a separate part for rectifying inspection is a great improvement. Additional material includes more discussion on Dodge's plans, with variation, similar schemes by Wald & Wolfowitz and Girshick, a brief survey of recent U.S. Government handbooks, and a discussion of Anscombe's scheme (12 pages).

Part IV. Control charts. The extra material in this part is mainly a discussion on the design of continuous sampling plans with a summary of a theoretical study by Duncan based on minimizing over-all cost, and a short summary of other control charts not given in detail (11 pages).

Part V. Some statistics useful in industrial research. About two-thirds of the extra material is in this section. The analysis of variance alone has 70 extra pages, including general linear comparisons, nested classifications, subgrouping and ranking means, confidence intervals and tables of expected values of mean squares for various models of components of variance analysis. A new chapter on the analysis of covariance is very clear, including both theory and numerical examples (33 pages). An extra chapter on the design of experiments deals with confounding and fractional replication (14 pages). Lastly, there is a new chapter on mapping response surfaces, with a detailed description of the Box–Wilson technique (44 pages). The chapter is typical of the book, with a clear presentation of the theory illustrated with numerical examples. The additions to this last part of the book represent a considerable improvement on the previous edition.

Many of Cox's criticisms of the first edition still stand. The author has obviously considered and anticipated such criticisms, for in his preface he says:

'As it now stands, the text is reasonably comprehensive and integrated. It nevertheless includes very little on non-parametric tests, nothing on stratified and multistage sampling and a meagre amount of material on the economics of statistical procedures. Several excellent volumes are available, however, on stratified sampling and other aspects of sampling survey techniques, and with the recent appearance of two books devoted entirely to non-parametric methods, it was felt that neglect of this important subject was justified. Work on economics of statistical procedures is very new and does not yet appear to be sufficiently standardized for comprehensive treatment in a text.'

In this paragraph the author himself gives the main criticism liable to arise, and we may well not agree with his defence for omitting the important material to which he refers. The reviewer feels that for a reference book of this type, some small space at least could have been given to sampling techniques, and two or three of the most useful non-parametric techniques, such as the Fisher–Terry normalized ranks, and the Mann–Whitney and Kolmogorov–Smirnov tests would have added little extra to the length. Indeed, one could well argue that the 127 page introduction describing probability, frequency distributions, etc., would be better spent on such material.

G. B. WETHERILL

Health and Medical Care in New York. Report by Committee for the Special Research Project in the Health Insurance Plan of Greater New York. Harvard: University Press for Commonwealth Fund; London: Oxford University Press. 1957. Pp. 275. 60*s*.

The Health Insurance Plan for New York is an organization which provides, on a pre-paid insurance basis, medical services for large groups of workers and their families. The doctors concerned work in partnerships or group practices and are paid a capitation fee. The dearth of actuarial data in the United

States on the demands and costs of such a service led to this survey of the working of the service using both the routinely kept records of the participating physicians and a sampling inquiry designed to elicit the medical needs and demands of both the families enrolled in the plan and of other families in New York. At the same time, an attempt was made to estimate the quality of medical care received by both population groups, either from the Health Insurance Plan and from other sources. There was also some interest in the possibility of using the H.I.P. experience as a guide to the medical service needs of the whole New York population.

The sampling methods used are discussed in some detail. From the total H.I.P. population, a 10% quasi-random sample was drawn by selecting each family whose enrolment certificate number ended in 5. The remainder of the New York population was sampled on an area probability basis; from 1000 city blocks, divided into fifty strata each of twenty blocks according to population characteristics such as density and recent increase, were drawn a proportion of families to give a total sample of 5000. The practical difficulties encountered in carrying out this plan are very fully and frankly discussed and the derivation of sampling errors for the results eventually obtained is clearly set out.

The design of the questionnaire used in the field survey owes much to the wide experience in this field which has been gathered by the U.S. Bureau of the Census and the detailed instructions to interviewers on the interpretation and coding of alternative answers form a most useful model for others venturing into this type of medical survey. Tabulation followed a prearranged scheme designed to answer the main questions posed by the guiding committee and to provide appropriate cross-checks on the accuracy of the information obtained. The preliminary analyses showed that there were severe practical limitations on the accuracy of certain types of data, such as income, given by the single informant for each family. On the other hand, comparisons between some of the results obtained and other official data showed a satisfying consistency.

The information on the prevalence and incidence of acute and chronic disease has its own intrinsic interest and the differences in their use of medical facilities between the samples is most interestingly discussed. The emphasis in this discussion is administrative rather than clinical: hospital admissions and frequency of dental treatment in relation to the educational status of the head of the household are typical subjects of particular interest. As an extremely competent description of a medico-social survey this book is to be commended. Lavishly produced, 275 pages long with much tabulated material, it may have a limited audience, but it should be included in any specialized library of reference texts on medical or social surveys.

D. D. REID

Tuberculosis in White and Negro Children. Harvard: University Press for Commonwealth Fund; London: Oxford University Press. 1958. **1. The Roentgenologic Aspects of the Harriet Lane Study.** By JANET B. HARDY. Pp. 122. 60*s*. **2. The Epidemiological Aspects of the Harriet Lane Study.** By MIRIAM E. BRAILEY. Pp. 103. 36*s*.

This study of tuberculosis in childhood is published in two volumes and is based on the wide experience of the disease among 1329 children, 437 from 180 white families, and 892 from 327 negro families. These were patients in the special Tuberculosis Clinic of the Harriet Lane Home of the Johns Hopkins Hospital in Baltimore. This Home was established in 1928 with such a study of the natural history of tuberculosis in children in mind and it is clear from these volumes that an immense effort has gone into the realization of this scheme. On the other hand, it should be pointed out at the onset that the area served was one of the poorest in Baltimore and both white and negro families lived in crowded and often insanitary homes. Again, the period covered, from 1928 to 1950, preceded the development and wide use of the powerful new antibiotics against tuberculosis which have radically altered the outlook in this disease. For these reasons, some at least of the findings cannot be directly applied to other communities or to the present time. Nevertheless, there is much of value in these volumes that will repay attention.

From the viewpoint of the biometrician new to the field of tuberculosis the first volume will provide a very clear introduction to the clinical concepts and terms in current use. Dr Janet B. Hardy, the present Director of the Clinic, gives a commentary on an excellent series of chest roentogenograms of a wide range of tuberculous conditions pointing out the characteristics important in their interpretation and classification. The second section of Volume I discusses the indications and techniques of bronchoscopy and bronchography.

Volume II, written by a previous Director, Dr Miriam E. Brailey, deals with the epidemiological aspects of the study. Since this work was done in collaboration with the Department of Epidemiology at Johns Hopkins, it necessarily reflects the concepts and technical methods developed by Wade Hampton Frost and his successors. This means a careful and critical attention to detail and the use of life-table techniques in the analysis of survival rates and reinfection rates in varying circumstances. There is a good discussion of the relevance of the findings to the vexed problem of the hazards of superinfection and although some may dispute a few of the conclusions this text is to be recommended to any serious student of the complexities of the epidemiology of tuberculosis.

D. D. REID

The Structure of Arithmetic and Algebra. By M. H. MARIA. New York: John Wiley and Sons Inc.; London: Chapman and Hall Ltd. 1958. Pp. vii + 294. 48*s*.

In 1547 William Buckley, Englishman, wrote a treatise entitled *The Rules of Arithmetic*, which [treatise] is noted chiefly for a short discussion on combinations. Much of the book, however, is concerned with rules for writing down numbers, for addition and subtraction and the other arithmetic operations. An expatriate from Mars who left England in 1547 might accordingly be predisposed to wonder at the lack of progress of English science since Prof. Maria is concerned to discuss many of the same topics as Buckley. Prof. Maria is, in fact, the latest of a long line of mathematical logicians who are concerned to describe on what mathematical thought is based. She does not waste much space on what is a number but moves briskly through the operations of addition, subtraction, multiplication and division. There are chapters on rational numbers, and actual numbers, irrational numbers and the properties of integers. In the generalization of the distributive law a little combinatorial theory is introduced.

This book may be interesting to the beginner in mathematical lore if he is prepared to recognize that in the end he will have to take almost everything here discussed for granted. For the statistician, who in many circumstances would be hard pressed to defend the logic of his operations, it is unlikely that this book will have much appeal.

F. N. DAVID

Tables of the Normal Probability Integral. (Sandia Corporation Technical Memorandum.) By D. T. MONK and D. B. OWEN. U.S. Department of Commerce: Office of Technical Services. 1957. Pp. 58. 40 cents.

The function $$G(h) = \frac{1}{\sqrt{(2\pi)}} \int_{-\infty}^{h} e^{-\frac{1}{2}x^2}\, dx$$

is tabulated to eight decimal places for $h = 0{\cdot}000\,(0{\cdot}001)\,7{\cdot}000$. An auxiliary table of Mills' Ratio

$$M(h) = \frac{1}{\sqrt{(2\pi)}} \int_{h}^{\infty} e^{-\frac{1}{2}x^2}\, dx \Big/ \frac{1}{\sqrt{(2\pi)}} e^{-\frac{1}{2}h^2}$$

is given to ten decimal places in $M(h)$ and for $h = 50\,(1)\,500$.

The Bivariate Normal Probability Distribution. (Sandia Corporation Research Report.) By D. B. OWEN. U.S. Department of Commerce: Office of Technical Services. 1957. Pp. 136. 65 cents.

Various methods have been used to calculate bivariate normal probabilities where the boundaries of the region over which the probabilities are to be computed are straight lines. This publication is mainly concerned with the use and tabulation of an auxiliary function $T(h, a)$ defined below. If

$$f(x, y; \rho) = \frac{1}{2\pi\sqrt{(1-\rho^2)}} \exp\left[-\frac{1}{2(1-\rho^2)}(x^2 - 2\rho xy + y^2)\right]$$

and $$G(h) = \frac{1}{\sqrt{(2\pi)}} \int_{-\infty}^{h} e^{-\frac{1}{2}t^2} dt,$$

then $$B(h, k; \rho) = \int_{-\infty}^{h} \int_{-\infty}^{k} f(x, y; \rho)\, dx\, dy$$

can be put in terms of $G(h)$, $G(k)$ and a function

$$T(h, a) = \frac{1}{2\pi} \int_{0}^{a} \frac{e^{-\frac{1}{2}h^2(1+x^2)}}{1+x^2} dx.$$

In a 76 pp. Appendix, $T(h, a)$ is tabulated to six decimal places for

$$a/h = 0{\cdot}000\,(0{\cdot}025)\,1{\cdot}000 \quad \text{and} \quad h = 0{\cdot}00\,(0{\cdot}01)\,4{\cdot}75.$$

Tables of the Bivariate Normal Distribution Function and Related Functions. Computed and collated under the direction of GERTRUDE BLANCH. U.S. Department of Commerce: National Bureau of Standards. Applied Mathematics Series 50. 1959. Pp. lxv + 258. $3·25

This is a definitive table of bivariate normal distribution, amplifying and extending the original tables computed under Karl Pearson's direction, which were brought together in *Tables for Statisticians and Biometricians*, 2, now out of print. Write

$$L(h, k, \rho) = \int_{h}^{\infty} dx \int_{k}^{\infty} g(x, y, \rho)\, dy,$$

where $$g(x, y, \rho) = \frac{1}{2\sqrt{(1-\rho^2)}} \exp\left[-\tfrac{1}{2}(x^2 - 2\rho xy + y^2)/(1-\rho^2)\right]$$

and $$V(h, k) = \frac{1}{2\pi} \int_{0}^{h} e^{-\frac{1}{2}x^2} dx \int_{0}^{kx/h} e^{-\frac{1}{2}y^2} dy.$$

Then the volume contains the following tables in addition to a 45 pp. Introduction.

Table I. Six decimal places given: $L(h, k, \rho)$ for $h, k = 0\,(0{\cdot}1)\,4$; $\rho = 0\,(0{\cdot}05)\,0{\cdot}95\,(0{\cdot}01)\,1$.

Table II. Seven decimal places given: $L(h, k, -\rho)$ for $h, k = 0\,(0{\cdot}1)\,h_n, k_n$; $\rho = 0\,(0{\cdot}05)\,0{\cdot}95\,(0{\cdot}01)\,1$, where $L(h_n, k_n, -\rho) \leqslant \frac{1}{2} \times 10^{-7}$ if h_n and k_n are both less than 4.

Table III. Seven decimal places given: $V(h, \lambda h)$ for $\lambda = 0{\cdot}1\,(0{\cdot}1)\,1$; $h = 0\,(0{\cdot}01)\,4\,(0{\cdot}02)\,4{\cdot}6\,(0{\cdot}1)\,5{\cdot}6$ and ∞.

Table IV. Seven decimal places given: $V(\lambda h, h)$ for $\lambda = 0{\cdot}1\,(0{\cdot}1)\,1$; $h = 0\,(0{\cdot}01)\,4\,(0{\cdot}02)\,5{\cdot}6$ and ∞.

Table V. Eight decimal places given: $y = (\arcsin r)/(2\pi)$; $r = 0\,(0{\cdot}01)\,1$ (second central differences).

Tables of natural logarithms for arguments between five and ten to sixteen decimal places. U.S. Department of Commerce: National Bureau of Standards Applied Mathematics 53. 1958. (Supersedes Mathematical Table 12). Pp. xiii + 506. $4.

This table is a reissue of Volume 4 of a four-volume table of logarithms published in 1941. $\mathrm{Log}_e x$ is tabulated for $x = 5{\cdot}0000\,(0{\cdot}0001)\,10{\cdot}0000$ to sixteen decimal places. No differences are given.

The companion volume to this, $\log_e x$ for $x = 0{\cdot}0000\,(0{\cdot}0001)\,5{\cdot}0000$ was noticed in *Biometrika*, **41**, 286.

Tables and Percentage Points of the Distribution Function of a Product. By LEO A. AROIAN. California: Hughes Aircraft Company. 1957. Pp. ix + 99.

x and y are two normal uncorrelated variables with means m_1 and m_2 and variances σ_1^2 and σ_2^2, respectively. $\delta_i = m_i/\sigma_i$ $(i = 1, 2)$. (i) The 100α percentage points of the distribution of $z = xy/(\sigma_1\sigma_2)$ are tabulated for different values of δ_i $(i = 1, 2)$. (ii) The cumulative distribution function of z is also tabulated.

For (i), $\alpha = 0{\cdot}001, 0{\cdot}005, 0{\cdot}010, 0{\cdot}025, 0{\cdot}050, 0{\cdot}100, 0{\cdot}250, 0{\cdot}500, 0{\cdot}750, 0{\cdot}900, 0{\cdot}950, 0{\cdot}975, 0{\cdot}990, 0{\cdot}995, 0{\cdot}999$ and $\delta_1, \delta_2 = (0{\cdot}4)\,4{\cdot}0, 6{\cdot}0, 12{\cdot}0$. Four decimal places are tabulated.

For (ii) $\delta_1, \delta_2 = 0\,(0{\cdot}4)\,4{\cdot}0, 6{\cdot}0, 12{\cdot}0$ and $F(z) = F(xy)$ is tabulated against z to six decimal places. The argument of z varies, but in the middle range between $z = -4{\cdot}0$ and $4{\cdot}0$ the argument interval is $0{\cdot}2$ or $0{\cdot}1$.

Rectangular-Polar Conversion Table. (Royal Society Mathematical Tables, Vol. 2.) By E. H. NEVILLE. Cambridge University Press for the Royal Society. 1956. Pp. xxii + 109. 30*s*.

Given $x = r\cos\theta$ and $y = r\sin\theta$, this table is concerned with the evaluation of r and of θ (in degrees) and $\log r$ and of θ (in radians) for different values of x and y.

For $x = 1(1)105$ and $y = x(1)105$, $\log_e r$ is tabulated to 15 decimal places and θ is given in radians to 15 decimal places.

For $x = 1(1)105$ and $y = 1(1)x$, r is tabulated to 13 decimal places and θ is given in degrees to 13 decimal places.

Auxiliary tables are the natural logarithms of the integers 1(1)160 to 15 decimal places, and integral multiples of $\log_e 10$ for 1(1)20. Tables to assist in interpolation are also appended.

Corrigenda

In the Review section of the last issue of *Biometrika* (1959), **46**, 494, the price of *Wahrscheinlichkeitsrechnung und Mathematische Statistik*, by M. Fisz (Berlin: VED Deutscher Verlag der Wissenschaften) was quoted wrongly as DM. 14; the correct price is DM. 36.

Other books received

Sampling Inspection Tables: Single and Double Sampling. Second edition. By H. F. DODGE and H. G. ROMIG. New York: John Wiley and Sons Inc.; London: Chapman and Hall Ltd. 1959. Pp. 224. 64*s*.

Modern Statistical Methods: Descriptive and Inductive. By P. O. JOHNSON and R. W. B. JACKSON. Chicago: Rand McNally and Company. 1959. Pp. 514. $8.00.

Introduction to Probability and Statistics. By G. W. MCELGRATH and B. W. LINDGREN. New York and London: The Macmillan Company, New York. 1959. Pp. 277. 44*s*.

Individual Choice Behaviour. By R. DUNCAN LUCE. New York: John Wiley and Sons Inc.; London: Chapman and Hall Ltd. 1959. Pp. 153. 48*s*.

Dynamic Management Decision Games. By J. GREENE and R. SISSON. New York: John Wiley and Sons Inc.; London: Chapman and Hall Ltd. 1959. Pp. 84. 24*s*.

Banca y Seguros (Revista doctrinal de la Dirección General de Banca Bolsa e Inversiones. Num 1.) Madrid: Gabreiel de Usera. 1959. Pp. 197. Subscription 150 pesetas.

The Theory of Storage. By P. A. P. MORAN. London: Methuen and Co. Ltd.; New York: John Wiley and Co. 1959. Pp. 111. 13*s*. 6*d*.

The Theory of Matrices. Vols. 1 and 2. By F. R. GANTMACHER. New York: Chelsea Publishing Co. 1959. Pp. 374 and 276. $6.00 each.

Darwin's Biological Work. Some Aspects Reconsidered. By BELL, CHALLINOR, HALDANE, MARLER, WHITEHOUSE, and WILKIE. London: Cambridge University Press. 1959. Pp. 343. 40*s*.

String Figures. By W. W. R. BALL. New York: Chelsea Publishing Company. 1960. Pp. 136. $3.95.

Regression Analysis. By E. J. WILLIAMS. London: Chapman and Hall; New York: John Wiley and Sons Inc. 1959. Pp. 214. 60*s*.

Mathematical Statistics. By J. N. KAPUR and H. C. SAXENA. Delhi: S. Chand and Co. 1960. Pp. 400. Rs. 10.

Russian Review of Biology. Translated from Vol. 48, No. 1. London and Edinburgh: Oliver and Boyd. 1959. Pp. 110. £9 or $30 for the first nine issues; thereafter, £6. 10*s*. or $20 annual subscription.

Rapid Statistical Calculations. By M. H. QUENOUILLE. London: Charles Griffin and Co. Ltd. 1959. 10*s*.

Annuario Estadistico de Espana for year 1959. Madrid: Instituto Nacional de Estadistica. 1959. Pp. xxxviii + 1099.

BAC

These
packin

Chequ
THE
STRI
pages

BIOMETRIKA

BACK ISSUES

Volumes 42–46 (1955–1959)

These may be obtained from the BIOMETRIKA OFFICE, at the following prices, which include packing and postage:

£3. 10*s*. (or $10.00) per wrappered volume.
Bound volumes: £1. 8*s*. (or $4.00) extra per volume.
Binding cases: 14*s*. (or $2.00) each.

Cheques should be made payable to *Biometrika*, crossed 'a/c BIOMETRIKA TRUST' and sent to THE SECRETARY, BIOMETRIKA OFFICE, UNIVERSITY COLLEGE, GOWER STREET, LONDON, W.C. 1, to whom all orders for subscriptions, offprints, and table separates (see pages (iv) , (v) and (vi)) should also be sent.

Volumes 20A, 20B, 21–41 (1928–1954)

Messrs Wm. Dawson & Sons, Ltd., have been appointed agents for the sale of these volumes, which may be obtained from the address below at the following prices, *packing and postage extra:*

£5. 5*s*. (or $15.00) per wrappered volume.
Bound volumes: £1. 8*s*. (or $4.00) extra per volume.
Binding cases: 14*s*. (or $2.00) each.
Index (Subjects Vols. 1–37; Authors Vols. 1–43): 6*s*. (or $1.00).

Volumes 1–19 (1902–1927)

Messrs Wm. Dawson & Sons, Ltd., have been granted permission to reprint Vols. 1–19. These are ready for distribution, price £130 bound. Single vols. £7. 10*s*. each. Would librarians and others wishing to have copies please place their orders with:

WM. DAWSON & SONS, LTD.,
16 WEST STREET, FARNHAM, SURREY, ENGLAND

(*All rights reserved*)

BIOMETRIKA, Vol. 47, Parts 1 and 2

CONTENTS

PAGE

M. S. Bartlett, J. C. Gower and P. H. Leslie. A comparison of theoretical and empirical results for some stochastic population models 1
David G. Kendall. Birth-and-death processes, and the theory of carcinogenesis 13
J. R. Ashford, C. S. Smith and Susannah Brown. The quantal response analysis of a series of biological assays on the same subjects 23
A. M. Walker. Some consequences of superimposed error in time series analysis 33
P. D. Finch. Deterministic customer impatience in the queuing system GI/M/1 45
D. E. Barton, F. N. David and Evelyn Fix. The polykays of the natural numbers 53
A. B. L. Srivastava. The distribution of regression coefficients in samples from bivariate non-normal populations. I. Theoretical investigation 61
Marvin A. Kastenbaum. The separation of molecular compounds by countercurrent dialysis: a stochastic process 69
J. G. Saw. A note on the error after a number of terms of the David–Johnson series for the expected values of normal order statistics 79
G. S. Watson. More significance tests on the sphere 87
N. L. Johnson. An approximation to the multinomial distribution: some properties and applications 93
M. K. Vagholkar and G. B. Wetherill. The most economical binomial sequential probability ratio test 103
J. R. McGregor. An approximate test for serial correlation in polynomial regression 111
Joseph Berkson. Nomograms for fitting the logistic function by maximum likelihood 121
Frank A. Haight and Melvin Allen Breuer. The Borel–Tanner distribution 143
E. J. Burr. The distribution of Kendall's score S for a pair of tied rankings 151

Miscellanea

M. E. Wise. On normalizing the incomplete beta-function for fitting to dose-response curves . . 173
R. G. Laha and E. Lukacs. On certain functions of normal variates which are uncorrelated of a higher order 175
H. D. Patterson. A further note on a simple method for fitting an exponential curve . . 177
Colin R. Blyth and G. L. Curme. Estimation of a parameter in the classical occupancy problem 180
C. W. Clunies-Ross and R. H. Riffenburgh. Geometry and linear discrimination 185
D. H. Bhate. Appproximation to the distribution of the sample size for sequential tests. II. Tests of composite hypotheses 190
P. D. Finch. On the covariance determinants of moving-average and autoregressive models . 194
P. H. Leslie. A note on some approximations to the variance in discrete-time stochastic models for biological systems 196
Coleridge A. Wilkins. On two queues in parallel 198
G. L. Turin. The characteristic function of Hermitian quadratic forms in complex normal variables 199
C. Burrows. Some numerical results for waiting times in the queue $E_k/M/1$ 202
B. Raja Rao. A formula for the curvature of the likelihood surface of a sample drawn from a distribution admitting sufficient statistics 203

Reviews

E. J. Gumbel. Statistics of Extremes 209
International Journal of Abstracts: Statistical Theory and Method, Vol. 1, No. 1 (Ed. W. R. Buckland) 209
E. L. Lehmann. Testing Statistical Hypothesis 210
J. Pfanzagl. Die axiomatischen Grundlagen einer allgemeinen Theorie des Messens 210
Gunnar Blom. Statistical Estimates and Transformed Beta-Variables 210
D. A. S. Fraser. Statistics: An Introduction 211
James H. Batchelor. Operations Research. An Annotated Bibliography. Second edition . . 211
Acheson J. Duncan. Quality Control and Industrial Statistics. (Revised edition) 212
Report by Committee for the Special Research Project in the Health Insurance Plan of Greater New York. Health and Medical Care in New York 212
Tuberculosis in White and Negro Children:
1. Janet B. Hardy. Aspects of the Harriet Lane Study 213
2. Miriam E. Brailey. Epidemiological Aspects of the Harriet Lane Study 214
M. H. Maria. The Structure of Arithmetic and Algebra 214
D. T. Monk and D. B. Owen. Tables of the Normal Probability Integral 214
D. B. Owen. The Bivariate Normal Probability Distribution 214
National Bureau of Standards:
Applied Mathematics Series 50. Tables of the Bivariate Normal Distribution and Related Functions 215
Applied Mathematics Series 53. Tables of natural logarithms for arguments between five and ten to sixteen decimal places 215
Leo A. Aroian. Tables and Percentage Points of the Distribution Function of a Product . . 215
E. H. Neville. Rectangular-Polar Conversion Table (Royal Society Mathematical Tables, Vol. 2) 216

Corrigenda 216

Other Books Received 217

Printed in Great Britain at the University Press, Cambridge (Brooke Crutchley, University Printer)